THE JOURNAL OF FINANCE • VOL. LIX, NO. 3 • JUNE 2004

Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards

WERNER ANTWEILER and MURRAY Z. FRANK*

ABSTRACT

Financial press reports claim that Internet stock message boards can move markets. We study the effect of more than 1.5 million messages posted on Yahoo! Finance and Raging Bull about the 45 companies in the Dow Jones Industrial Average and the Dow Jones Internet Index. Bullishness is measured using computational linguistics methods. *Wall Street Journal* news stories are used as controls. We find that stock messages help predict market volatility. Their effect on stock returns is statistically significant but economically small. Consistent with Harris and Raviv (1993), disagreement among the posted messages is associated with increased trading volume.

MANY PEOPLE ARE DEVOTING a considerable amount of time and effort creating and reading the messages posted on Internet stock message boards. News stories report that the message boards are having a significant impact on financial markets. The Securities and Exchange Commission has prosecuted people for Internet messages. All this attention to Internet stock messages caused us to wonder whether these messages actually contain financially relevant information.[1] We consider three specific issues. Does the number of messages posted or the bullishness of these messages help to predict returns? Is disagreement among the messages associated with more trades? Does the level of message posting or the bullishness of the messages help to predict volatility?

The first issue is, does the level of message activity or the bullishness of the messages successfully predict subsequent stock returns? This is the natural starting place because a very high proportion of the messages contain explicit assertions that the particular stock is either a good buy or a bad buy. Of course, there are a great many previous empirical studies showing how hard it is to predict stock returns by enough to cover transactions costs. We find that there is evidence of a small degree of negative predictability even after controlling for bid–ask bounce. When many messages are posted on a given day, there

*Both authors are at the Sauder School of Business, University of British Columbia. We would like to thank Richard Arnott; Jack Cooney; Elizabeth Demers; Rick Green; Alan Kraus; John Ries; Jacob Sagi; an anonymous referee; and the seminar audiences at UBC, the University of Toronto, the 2002 Canadian Economics Association, the Nasdaq-Yale-JFM conference, the 2002 Northern Finance Association, and the 2002 American Finance Association annual meetings for helpful comments. We thank the SSHRC (501-2001-0034) for financial support. The second author thanks the B. I. Ghert Family Foundation for financial support. All errors are ours.

[1] Weiss (2000) provides a nice discussion of how message boards were regarded in the financial press as of 2000.

is a statistically significant negative return on the next day. This return is economically very small in comparison to plausible transactions costs.

The second issue is whether greater disagreement is associated with more trades. This question was stimulated by reading the messages. At times the message boards reflect considerable disagreement, while at other times much greater consensus is evident.

Financial theory provides two distinct perspectives on disagreement. A traditional hypothesis is that disagreement induces trading. Hirshleifer (1977), Diamond and Verrecchia (1981), Karpoff (1986), Harris and Raviv (1993) and others have carried out related theoretical analysis. An alternative perspective is found in the "no-trade theorem" of Milgrom and Stokey (1982). According to the "no-trade theorem," when one person considers trading with another person, each of them needs to consider why the other person might be willing to trade at a particular price. In some settings disagreement does not induce trading; it leads to a revision of market prices and beliefs.

We consider differences of opinion both in contemporaneous regressions and predictive regressions. We also condition on a variety of other factors. In the contemporaneous regressions disagreement is associated with more trades. However, there is a reversal on the next day such that trading volume is lower than it otherwise would have been.

The third issue is whether the message boards help to predict volatility. A remarkable range of sometimes quite odd things are said in the messages. This leads to the hypothesis that perhaps the people posting the messages are real-world counterparts of the "noise traders" that are so often invoked in financial theory. In order to test this idea we need to define and to model volatility. The literature provides a large number of approaches to volatility.

Many recent studies have employed realized volatility instead of focusing on the squared residuals from a returns regression.[2] Andersen et al. (2001) have used this approach in a study of the firms in the Dow Jones Industrial Average. We consider the importance of the message boards within a fractionally integrated realized volatility news response function that follows their approach. We also consider volatility vector autoregression that is related to Andersen et al. (2002).

The GARCH class of volatility models remains popular. Recent studies such as Hansen and Lunde (2001) and Engle and Patton (2001) show that it is hard to beat a GARCH (1,1) within the class of GARCH models. However, there is some evidence of an asymmetric response between positive and negative shocks, as in Glosten, Jagannathan, and Runkle (1993). Accordingly, we have also considered the effect of the message boards within the context of GARCH, EGARCH, and GJR models. To save space these results are left to the technical appendix.[3]

Volatility models are often estimated without exogenous variables. However, it is well known that trading volume helps forecast volatility (see Jones, Kaul,

[2] Realized volatility follows from the work of French, Schwert, and Stambaugh (1987) and Schwert (1990). It has been given theoretical foundations by Andersen et al. (2001).

[3] A technical appendix to this paper is available on the web as a PDF file at http://pacific.commerce.ubc.ca/antweiler/public/noise-1.pdf. The appendix contains additional discussions, results, and robustness checks.

and Lipson (1994) for example). Glosten et al. (1993) and Engle and Patton (2001) also fit models that include the Treasury bill rate as an exogenous factor. We found evidence supporting the role of trading volume, but for our sample we did not find any evidence that the Treasury bill rate helped to forecast volatility. Thus, we include trading volume as an added factor in our volatility tests.

We find that message posting helps to predict volatility. Perhaps due to multicollinearity, adding trading volume tends to reduce the impact of the number of messages on market volatility. However, this reduction does not cause the message board effect to vanish. Trading volume is the more important factor for predicting the market volatility of some firms, while the messages are more important for predicting the market volatility of other firms. Evidence for an effect of bullishness or disagreement on volatility is weak.

Why do people post messages on Internet stock message boards? To properly answer this question requires a theory of communication that also contains a financial market. Such theories are starting to be developed. DeMarzo, Vayanos, and Zwiebel (2001) argue that people overweight the opinions of those with whom they talk. This kind of belief formation process can make it profitable to be an influential agent. In the equilibrium all agents will want to listen to other agents who are particularly influential since what they say will affect the market. The model is of particular interest for our purposes since it provides an explanation both for why people post messages on message boards, and for why other people might choose to read the message boards.[4]

A somewhat different perspective is offered by Cao et al. (2002). They model the importance of fixed costs of market participation, which implies that potential traders do not always trade. They argue that conversation is then potentially important: "The introduction of conversation among a subset of market participants may have large effects on the equilibrium. A sidelined investor who learns that another investor shares a similar signal may decide to participate. (p. 644)" If the stock message boards permit this type of communication, then the prediction is that message posting should be followed by trading. Our evidence supports this prediction.

The previous literature contains a small number of papers that have examined the ability of message boards to predict stock returns. There are mixed claims about whether public information on the Internet can predict subsequent stock returns. Our finding that higher message postings predict negative subsequent returns has not previously been reported. Our result does seem to be economically small but statistically robust. The previous literature has not examined the issue of whether differences of opinion in the stock messages are associated with more trades. We find that differences of opinion are associated

[4] Shiller (2000) also draws attention to the role of conversation, and suggests that information passed through conversation may play an important role in informational cascades. Hong, Kubik, and Stein (2002) provide indirect evidence of the importance of word of mouth communication. They show that mutual fund manager's trades in a given stock are connected to the trading decisions of other fund managers located in the same city. They interpret the findings in terms of an epidemic model of information spread by word of mouth. In contrast to the current paper, they can observe the trader's portfolios, but they do not have direct measures of the communication.

with more trades. Similarly, previous studies have not examined the connection between stock message posting and stock market volatility.

Turning to the individual papers in the previous literature, the first study of Internet stock message boards was Wysocki (1999). For the 50 firms with the highest posting volume between January and August 1998, he reports that message posting did forecast next-day trading volume and next-day abnormal stock returns. Using a broader sample of firms, Wysocki also measured the cumulative message postings on Yahoo! Finance to July 1, 1998, and studied the cross-sectional differences between firms in order to determine which firms had a large number of messages posted. The firms with high message posting activity were characterized by high market valuation relative to fundamentals; high short seller activity; high trading volume; and high analyst following but low institutional holdings.

Bagnoli, Beneish, and Watts (1999) compared the First Call analyst earning forecasts to unofficial "whispers." The whispers were collected from a number of sources including Internet web pages and news stories that reported the whisper forecasts. The analysts from First Call tended to underestimate corporate earnings announcements, while the whispers tended to be more accurate.

In a study of stocks in the Internet service sector, Tumarkin and Whitelaw (2001) found that the messages did not predict industry adjusted returns or abnormal trading volume. Das and Chen (2001) is devoted to the development of a new natural language algorithm to classify stock messages. They illustrate its application on a selected sample of nine firms during the last quarter of 2000. They find that the stock messages reflect information rapidly but do not forecast stock returns. Dewally (2000) collected stock recommendations from two newsgroups (misc.invest.stocks and alt.invest.penny-stocks). He found that there was no predictive content in the forecasts on these newsgroups. The recommended stocks typically had strong prior performance.

The rest of the paper is organized as follows. Section I discusses the messages and how we extracted information from the texts. In Section II we describe a number of the basic features of the data. The predictability of stock returns is considered in Section III. Section IV presents the volatility results. Trading volume is considered in Section V. Both the effect of disagreement on trading volume and the predictability of trading volume are studied. The role of the *Wall Street Journal* is considered in Section VI. We conclude in Section VII.

I. Message Board Data and Classification

During 2000, Yahoo! Finance and Raging Bull provided two of the largest and most prominent sets of message boards. The sample of stocks being studied are the 45 firms that together made up the Dow Jones Industrial Average (DIA) and the Dow Jones Internet Commerce Index (XLK). These firms were fairly large and well known.

Messages were downloaded from the Yahoo! Finance (YF) and Raging Bull (RB) message boards using specialized software written by the authors. Downloaded messages were stored in a simple plain-text database format, one file

```
---------------------
FROM YF
COMP ETYS
MGID 13639
NAME CaptainLihai
LINK 1
DATE 2000/01/25 04:11
SKIP
TITL ETYS will surprise all pt II
SKIP
TEXT ETYS will surprise all when it drops to below 15$ a pop, and even then
TEXT it will be too expensive.
TEXT
TEXT If the DOJ report is real, there will definately be a backlash against
TEXT the stock. Watch your asses. Get out while you can.
---------------------
FROM YF
COMP IBM
MGID 43653
NAME plainfielder
LINK 1
DATE 2000/03/29 11:39
SKIP
TITL BUY ON DIPS - This is the opportunity
SKIP
TEXT to make $$$ when IBM will be going up again following this profit taking
TEXT bout by Abbey Cohen and her brokerage firm.
TEXT
TEXT IBM shall go up again after today.
-----------------
```

Figure 1. Samples of bulletin board messages.

per day per company. Each message is uniquely identified by the bulletin board code (YF or RB), the company's ticker symbol, and the message board sequence number. The file contents were then summarized in an index file that also lists the date and time of posting, the message's length in words, and the screen name of the originator of the message.

To understand the nature of the postings it is helpful to look at examples. Figure 1 provides two fairly typical examples of messages in the database format. Each message is dated and timed to the minute, has a title, and has a text. The text very often contains a predicted price change and at least some explanation for the prediction. Most of the explanations are fairly short. The number of words in a message is most frequently between 20 and 50. Relatively few messages have more than about 200 words. It is fairly rare for a message to have more than 500 words. More than 40% of the messages are posted by people who post only a single message.[5] However, there are some people who are very active and account for more than 50 messages each.

[5] We only observe the chosen screen name rather than the author's actual name. Therefore, if one author posts messages using more than one screen name we count these as if they were separate authors.

Our message boards data contain more than 1.5 million text messages—far too many to interpret manually. In order to assess the content of the stock messages we employ well established methods from computational linguistics. The oldest algorithm used to interpret text is called Naive Bayes. We use this classic algorithm as our main case. Another algorithm called Support Vector Machine has become very popular for use in many classification problems, including text classification.[6] In order to ensure robustness we carry out all tests using both algorithms to measure the messages. Both algorithms are used to code the individual messages as bullish, bearish, or neither. The two algorithms produce quite similar results and so we only report the Naive Bayes results in the text.

A. Naive Bayes Message Coding

The Naive Bayes algorithm is the oldest of the algorithms used to classify documents. Lewis (1998) provides a perspective on the history of the algorithm. It continues to be among the most successful natural language algorithms.

For Naive Bayes text classification we have employed the Rainbow package developed by McCallum (1996).[7] The key assumption underlying the Naive Bayes classification method is that occurrences of words are independent of each other. The assumption of independence among words is the reason that the algorithm is referred to as "naive." Even though this is a highly unrealistic assumption, Naive Bayes performs rather well in practice.[8]

In the context of text classification, Naive Bayes can be understood most easily as a straight-forward mechanism of updating odds ratios. Consider a stream of words W_i that are found either in a message of type T or its anti-type $\tilde{T}$. Let m be the number of occurrences of this word in type T, and let $\tilde{m}$ be the number of occurrences in anti-type $\tilde{T}$. Further, let n and $\tilde{n}$ denote the total number of words in classes T and $\tilde{T}$, respectively. For words found in messages from the training set we observe the conditional probabilities $P(W_i|T) = m_i/n_i$ and $P(W_i|\tilde{T}) = \tilde{m}_i/\tilde{n}_i$. Now consider Bayes' rule, updating our prior $P(T|W_{i-1})$ to posterior $P(T|W_i)$ when we observe word W_i and thus $P(W_i|T)$ and $P(W_i|\tilde{T})$:

$$P(T|W_i) = \frac{P(T|W_{i-1})P(W_i|T)}{P(T|W_{i-1})P(W_i|T) + (1 - P(T|W_{i-1}))P(W_i|\tilde{T})}. \tag{1}$$

[6] We discuss the Support Vector method in the technical appendix.

[7] This software can be downloaded freely for academic purposes from the web at http://www.cs.cmu.edu/ mccallum/bow/.

[8] This approach, an example of a "bag of words" approach to text classification, makes no direct use of the grammatical structure. As an empirical matter it has been found that a surprisingly small amount is gained at substantial cost by attempting to exploit grammatical structure in the algorithms. For a helpful discussion of the various approaches to analyzing text see Manning and Schütze (1999).

That is easily rewritten in odds-ratios form as

$$\frac{P(T|W_i)}{1 - P(T|W_i)} = \frac{P(T|W_{i-1})}{1 - P(T|W_{i-1})} \cdot \frac{P(W_i|T)}{P(W_i|\tilde{T})} \tag{2}$$

with $P(T|W_0) \equiv P(T)$. Classifying a document thus amounts to multiplying odds ratios when processing the document word by word. For reasons of computational accuracy, it is however common practice to add up logs of odds ratios,

$$P(T|W_N) = P(T) \exp\left[\sum_{i=1}^{N} \ln\left(\frac{P(W_i|T)}{P(W_i|\tilde{T})}\right)\right], \tag{3}$$

where N is the number of words in a given document.[9]

We start by manually classifying a training data set of 1,000 messages. Based on this training data set, the classification software filters our entire sample of 1,559,621 messages to obtain buy, hold, or sell signals for each message. We then aggregate the codings into indices that measure the bullishness of each stock message board during each time period. For our analysis, we study time periods of 15 minutes, one hour and one day.

Usage of the Rainbow software package proceeded in three steps. First we split the 1,000 messages into buy, sell, and hold messages stored in individual directories. In the second step we run the rainbow utility to process the messages in the training data set. In the software we use the option settings "naive bays" for method, and "1000" for prune-vocab-by-infogain. The latter restricts the number of words in the vocabulary to the top 1,000 words as ranked by the average mutual information with the class variable.[10] After training is complete, Rainbow is put into server mode and individual messages k containing words W^k are sent from a client program to the server for evaluation. The server returns three probabilities $P(c|W_N^k)$ for each of the three categories $c \in \{\text{BUY, HOLD, SELL}\}$, and we choose the classification with the highest probability according to $\arg\max_c P(c|W_N^k)$.

Table I provides a comparison between the manual classification of the messages and the classification by the Naive Bayes algorithm. While buy messages are five times more common than sell messages, two thirds of all messages are neither. The algorithm performs reasonably well, as is apparent from the relatively small number of misclassifications.

[9] Adding logs of odds ratios avoids the problem of computational "underflow" or "overflow" errors, which could easily arise when odds ratios are multiplied directly in a long message. The prior $P(T)$ is based on the document frequencies for each of our three classes. The problem with equation (1) is that either $P(W_i|T)$ or its anti-class counterpart may be zero. In this case, a method known as *Laplace smoothing* is applied to replace the zero-value with estimates $\mathcal{E}(P(W_i|T)) = (1 + m_i)/(1 + n_i)$, and likewise for $P(W_i|\tilde{T})$.

[10] See Cover and Thomas (1991) for details.

Table I
Naive Bayes Classification Accuracy within Sample and Overall Classification Distribution

The first percentage column shows the actual shares of 1,000 hand-coded messages that were classified as buy (B), hold (H), or sell (S). The buy-hold-sell matrix entries show the in-sample prediction accuracy of the classification algorithm with respect to the learned samples, which were classified by the authors (Us).

Classified:		By Algorithm		
by Us	%	Buy	Hold	Sell
Buy	25.2	18.1	7.1	0.0
Hold	69.3	3.4	65.9	0.0
Sell	5.5	0.2	1.2	4.1
1,000 messages[a]		21.7	74.2	4.1
All messages[b]		20.0	78.8	1.3

[a]These are the 1,000 messages contained in the training data set.
[b]This line provides summary statistics for the out-of-sample classification of all 1,559,621 messages.

B. Aggregation of the Coded Messages

We aggregate the message classifications x_i in order to obtain a bullishness signal B_t for each of our time intervals t. Let $M_t^c \equiv \sum_{i \in \mathcal{D}(t)} w_i x_i^c$ denote the weighted sum of messages of type $c \in \{\text{BUY}, \text{HOLD}, \text{SELL}\}$ in time interval $\mathcal{D}(t)$, where x_i^c is an indicator variable that is one when message i is of type c and zero otherwise, and w_i is the weight of the message. When the weights are all equal to one, M_t^c is simply the number of messages of type c in the given time interval. Furthermore, let $M_t \equiv M_t^{\text{BUY}} + M_t^{\text{SELL}}$ be the total number of "relevant" messages, and let $R_t \equiv M_t^{\text{BUY}}/M_t^{\text{SELL}}$ be the ratio of bullish to bearish messages.

There are many different ways that M_t^{BUY}, M_t^{HOLD}, and M_t^{SELL} can be aggregated into a single measure of "bullishness," which we denote as B. We have considered the empirical performance of a number of alternatives, and the results are generally quite similar. One way to think about the aggregation question is to ask, what is the degree of homogeneity that we should choose for the aggregation function? Our first measure is defined as

$$B_t \equiv \frac{M_t^{\text{BUY}} - M_t^{\text{SELL}}}{M_t^{\text{BUY}} + M_t^{\text{SELL}}} = \frac{R_t - 1}{R_t + 1}. \tag{4}$$

It is homogenous of degree zero and thus independent of the overall number of messages M_t. Multiplying M_t^{BUY} and M_t^{HOLD} by any constant will leave B_t unchanged. Furthermore, B_t is bounded by -1 and $+1$. We investigated this measure extensively in an earlier draft of this paper. While all our key results can be obtained with this measure, we prefer the following second approach to defining bullishness:

$$B_t^* \equiv \ln\left[\frac{1+M_t^{\mathrm{BUY}}}{1+M_t^{\mathrm{SELL}}}\right] = \ln\left[\frac{2+M_t(1+B_t)}{2+M_t(1-B_t)}\right] \approx B_t \ln(1+M_t). \tag{5}$$

This measure takes into account the number of traders expressing a particular sentiment. This aggregation function is homogenous of a degree between zero and one. We have also systematically investigated a third aggregation function

$$B_t^{**} \equiv M_t^{\mathrm{BUY}} - M_t^{\mathrm{SELL}} = M_t\left[\frac{R_t-1}{R_t+1}\right] = M_t B_t, \tag{6}$$

which is homogenous of degree one. The last two measures both increase with the number of messages M_t as well as the ratio of bullish to bearish messages. However, the logarithmic transformation in our "intermediate" measure B^* discounts excessively large ratios or message numbers. The measure B^* appears to outperform both alternatives and so we use it in all reported tables.[11]

In the reported tables we focus on the case in which each message is weighted equally ($w_i = 1$). In addition to equal weighting, we considered two unequal weighting schemes. Because longer messages might be more important than short ones, we tried weighting the messages by their length ($w_i = \mathrm{LENGTH}_i$). Some people post many very similar messages. Because repeated messages by the same author may have diminishing impact, we also tried weighting each message by the inverse of the total number of messages posted by the author of the particular message ($w_i = \mathrm{AUTHOR}_i^{-1}$). These alternative weighting schemes do not affect any of our conclusions, and so we stick with the simplicity of equal weighting.

We also considered using weights based on the citation frequency of individual messages. This information is available for messages on the Yahoo! Finance boards, where messages may contain a single citation of a previous message. We do not employ such a weighting system in our analysis because citation weights are only determined ex post. The number of citations that a particular message will receive is unknown at the time it is posted. This would cause serious problems for our time sequencing analysis. Another concern is that citations may not be genuine citations of an earlier message. This happens because some message board contributors find it more convenient to use an e-mail system's "reply" function instead of the "new message" function. As a result the number of citations can be a distorted measure.

The treatment of periods during which no messages are posted is a potential source of concern. If no new messages are posted, does this mean that everyone agrees with the last message? Or does it simply mean that nobody has anything

[11] In none of our three measures do we use the number of "hold" messages. This group contains both "noise" as well as neutral (hold) opinions. The amount of "noise" dominates. Thus the inclusion of M_t^{HOLD} would lead to noisier and perhaps distorted bullishness signals.

that they wish to say? In the reported results we assume that an absence of postings is a zero.[12]

Disagreement has long been considered a possible motivation for trading. Important examples of papers presenting this idea include Hirshleifer (1977), Karpoff (1986), Harris and Raviv (1993), and Diamond and Verrecchia (1981). In a standard rational expectations model, however, there can be "no trade" theorems such as in the analysis of Milgrom and Stokey (1982). Thus theory allows for more than one perspective on the role of disagreement in financial markets. The traditional perspective is often thought to be the more plausible empirical approach.[13]

The message boards are a natural place to test this hypothesis since disagreement is rather directly observable. To measure agreement we proceeded as follows. First we define $x_i \equiv x_i^{\mathrm{BUY}} - x_i^{\mathrm{SELL}} \in \{-1, +1\}$. Again all hold messages are ignored. The variance of B_i during time interval t corresponding to (4) can then be calculated as

$$\sigma_t^2 \equiv \frac{\sum_{i \in \mathcal{D}(t)} w_i (x_i - B_t)^2}{\sum_{i \in \mathcal{D}(t)} w_i} = \frac{\sum_i w_i x_i^2}{\sum_i w_i} - B_t^2 = 1 - B_t^2. \tag{7}$$

In the last simplification step, observe that $\forall i : x_i^2 = 1$, because x_i is either $+1$ or -1. This permits us to measure the prevailing level of agreement among message board contributors as an "agreement index,"

$$A_t \equiv 1 - \sqrt{1 - B_t^2} \in [0, 1]. \tag{8}$$

If disagreement produces trading, then A_t should be negatively correlated with measures of trading volume.

To illustrate the properties of the agreement index A, consider the case where there are three messages in a particular time interval. First, assume that all three are bullish ($x_i = +1$). Then it is easily verified that $A = 1$. The same result emerges when all three messages are bearish ($x_i = -1$). Second, consider the case where two messages are bullish and one is bearish. In this case $B = 1/3$, and therefore $A = 1 - \sqrt{1 - (1/3)^2} = 0.057$, so agreement is low. Third, if we have one buy, one sell, and one hold message, the agreement index will be zero. Finally, what happens to the agreement index if we do not observe any messages in a given time period? In these periods we assume that our bullishness index is neutral ($B = 0$), and therefore the agreement index is calculated as zero as well. Intuitively, periods without information may be viewed as latent disagreement.

[12] Another possibility is to maintain the previous value instead of using a value of zero during a period without messages. We did some experiments with such a procedure. These indices are biased in the bullish direction. The empirical results obtained using this approach are considerably noisier.

[13] Related empirical evidence is provided by Kandel and Pearson (1995) and Bessembinder, Chan, and Seguin (1996).

II. Data Description

A. Messages

More messages are posted on Yahoo! Finance than on Raging Bull. Firms that are listed on Nasdaq and included in the XLK index generate more messages than do firms listed on the NYSE and included in the DIA. Intel and Microsoft are the only firms in our sample that are listed on Nasdaq and included in DIA. Many messages are posted about both of these firms.

Figure 2 shows the weekly level of message posting over the full year 2000. News reports in the earlier part of 2000 suggested that posting activity was increasing at a dramatic rate, and news reports in the later parts of 2000 suggested that posting activity was falling off dramatically. Neither of these reports match what we observed for our sample of firms. There was some decline in activity during the late spring and the summer months, but otherwise message posting activity was reasonably stable over the year for our sample of firms. Within the trading week there is a very strong weekend effect. Many fewer messages are posted during weekends.

B. Financial Data

Financial data are from the TAQ database for the 45 stocks and for the exchange-traded fund that serves as our proxy for the market. We extracted

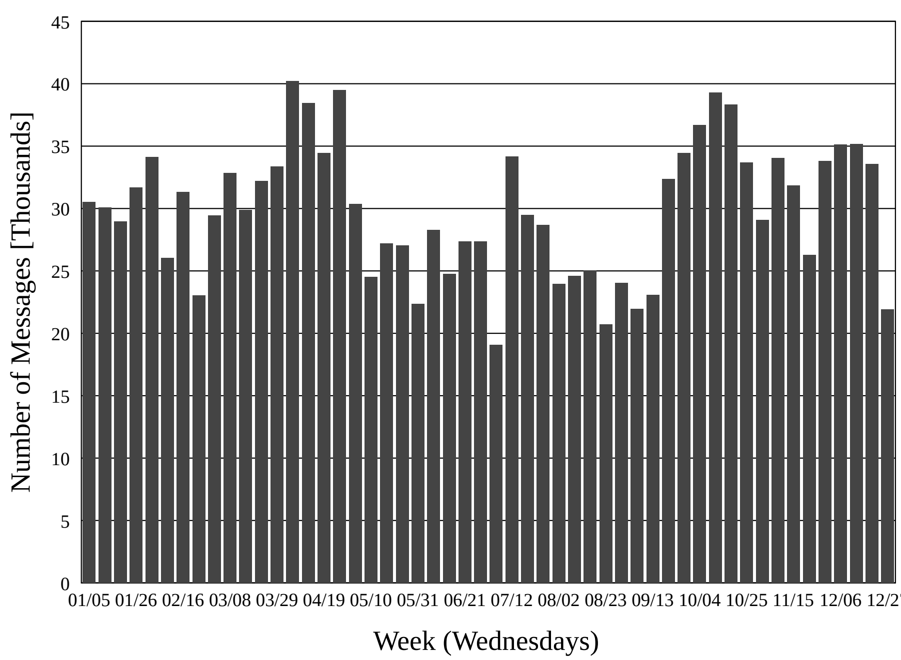

Figure 2. Weekly message board activity. Posting activity of 45 companies in DIA and XLK combined.

and then aggregated the following information: a bid–ask midpoint at the end of each 15-minute time interval; the corresponding bid–ask spread at that point; the volume-weighted average trading price for that 15-minute interval; the corresponding volume-weighted volatility; the number of shares traded; and the number of transactions in each of three transaction value categories (below \$100,000, between \$100,000 and \$1,000,000, and \$1,000,000 or above).

In scanning through the TAQ trades file we ignore information from exchanges other than Nasdaq or NYSE; we ignore trades or quotes with sequence numbers that are "out of sequence"; we ignore trades that are marked as irregular; and we ignore opening and closing quotes. In addition, we filter out spreads that are negative or in excess of 40% of the bid–ask midpoint, and we filter out trades with prices 100% above or below the bid–ask midpoint for the relevant time period.

Our measure of the stock market index is the exchange-traded fund that mimics the S&P 500 (ticker symbol SPY). This is a market factor that traders can buy and sell easily with low transaction costs. It has the further advantage that we can observe its market price directly at the same frequency as the rest of the financial data.

As can be seen in Figure 3, there was a decline in the volume of stock trading over the year for our sample of firms. Comparison of Figure 2 with Figure 3

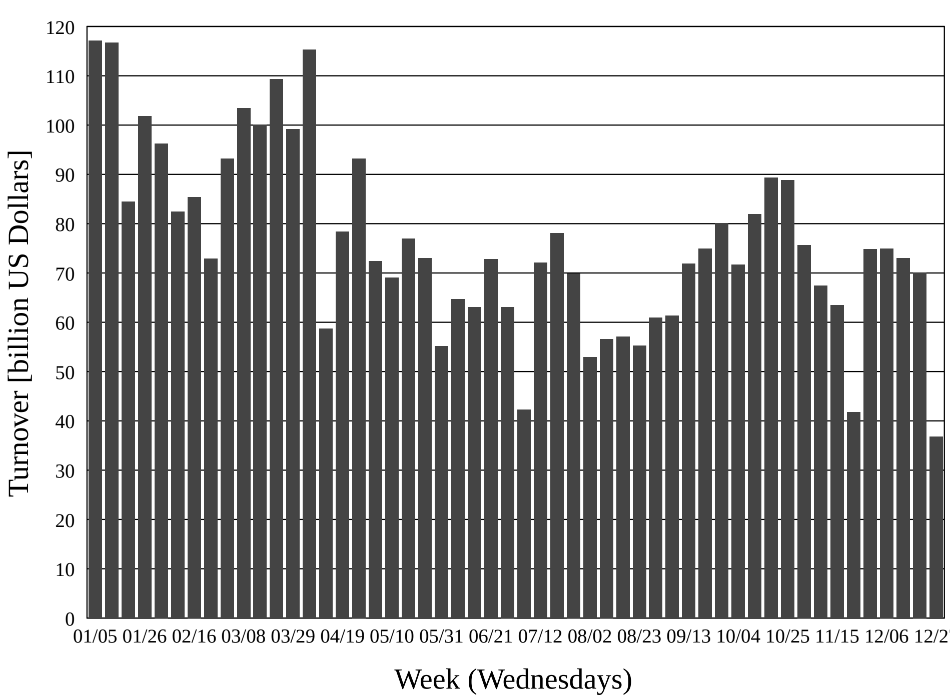

Figure 3. Weekly trading volume. Turnover of 45 companies in ECM and DJX combined.

suggests that trading volume is often elevated during the same weeks that message posting is elevated.

C. Wall Street Journal

Using Lexis/Nexis we collected all the articles about our 45 firms that appeared in the *Wall Street Journal* during 2000. As shown in Table II, there are no news stories about most firms on most days, while the occasional firm has more than a single news story on a given day. The main firm that had multiple news stories on the same date is Microsoft, which was involved in a controversial antitrust trial during the year. Table II also shows that while the message board postings are very heavy for the Internet firms, the *Wall Street Journal* (hereafter, WSJ) provides much more coverage for the typical Dow Jones Industrial Average firm. In other words, there is a difference in emphasis between these two sources.

The WSJ is usually taken to be the statement of record and it is published each business day morning. Investors, however, commonly obtain information from news wire sources directly or from a range of sources that post news on the Internet—including the web pages of the WSJ. This kind of information commonly becomes available during the day before it is published in the WSJ. Therefore, it is important to also examine the days surrounding the day that a story appears in the WSJ. We consider two days prior and one day subsequent to a news story day.

D. Time

The NYSE and Nasdaq are open from 9:30 a.m. to 4:00 p.m. Eastern time. As is well known, the behavior at the market open and at the close of trading have somewhat different properties than during the rest of the day. In particular, as shown in Figure 4, trading volume tends to be lower during the middle of the day. As explained by Bacidore and Lipson (2001), the opening and closing procedures on the NYSE are different from trading during most of the day. The opening auction can take up to half an hour.

In addition to the different trading institutions, there are also potentially different trader motivations to consider. Many small traders think about their portfolios during the evening when they are home from work. They may call their brokers or place automated trades through discount brokers before they start work in the morning. This will result in many trades at the market open. Near the end of the day institutional traders may wish to close out a position in order to avoid overnight risk. Managers of mutual funds may wish to prepare their portfolio for the end of the day valuation. These arguments are consistent with what is observed in Figure 4.

One might have imagined that a high proportion of the messages would be posted in the evening, after dinner. However, this is not the dominant pattern. Figure 5 shows that message posting is concentrated during working hours, while the stock markets are open. This is suggestive of day trader activity, but it could also reflect people posting messages from their jobs.

Table II

Summary Statistics by Company

The table is sorted by company group (DIA, XLK) and within each group by descending bullishness index averaged across the two bulletin boards.

	Bullishness[a]		Activity[b]		Intensity[c]		Return[d]	Vola.[e]	WSJ
Company Name	YF[f]	RB[f]	YF[f]	RB[f]	YF[f]	RB[f]	[%]	[–]	[#]
Philip Morris	0.597	0.325	78.4	5.8	74	115	86.5	4.22	45
Intel	0.632	0.300	80.2	14.1	52	85	−64.2	6.84	96
Microsoft	0.550	0.234	159.3	38.3	56	90	−63.0	5.00	397
General Electric	0.529	0.298	40.1	15.8	72	90	−68.7	4.47	96
AT&T	0.494	0.259	64.9	11.9	53	78	−66.5	4.59	189
Citigroup	0.251	0.407	4.4	2.5	60	97	−7.6	4.53	80
Wal Mart	0.309	0.334	20.5	3.2	82	79	−22.5	4.90	55
Hewlett Packard	0.313	0.238	16.0	0.8	63	125	−72.4	6.41	36
Honeywell	0.307	0.256	12.1	0.7	76	81	−18.3	5.10	27
Johnson&Johnson	0.302	0.298	2.9	0.4	70	72	12.7	2.97	31
Walt Disney	0.296	0.313	18.3	2.0	71	96	−0.9	4.29	83
Procter&Gamble	0.324	0.161	19.1	2.3	55	95	−27.2	3.58	54
Home Depot	0.265	0.372	17.6	3.0	54	81	−33.5	4.71	26
IBM	0.282	0.249	24.5	2.6	66	98	−24.6	4.46	108
SBC Communications	0.287	0.193	16.6	1.5	65	88	−1.9	4.07	44
United Technologies	0.259	0.289	2.0	0.1	79	70	20.9	4.66	18
Intn'l Paper	0.254	0.259	8.9	0.3	75	69	−28.4	5.08	19
Boeing	0.242	0.093	54.8	1.5	81	90	58.9	4.02	123
McDonalds	0.241	0.195	4.3	0.6	67	63	−15.0	3.96	32
Eastman Kodak	0.242	0.193	2.7	0.2	72	129	−41.1	3.67	18
JP Morgan	0.228	0.241	1.4	0.1	59	127	30.9	4.36	58
Alcoa	0.210	0.183	5.0	0.2	56	38	−59.3	5.22	26
American Express	0.201	0.284	3.4	0.2	58	70	−66.6	5.26	44
Minnesota Mining	0.184	0.287	1.5	0.2	66	101	25.6	3.63	6
Coca Cola	0.188	0.181	9.8	0.5	98	94	6.1	3.95	106
Du Pont	0.177	0.155	6.7	0.3	72	74	−26.8	4.44	0
Merck	0.169	0.131	8.8	0.5	75	87	36.8	3.45	8
Caterpillar	0.133	0.218	1.4	0.2	65	60	−3.8	4.51	8
General Motors	0.136	0.154	6.6	0.6	85	96	−31.3	3.88	181
Exxon	0.113	0.158	7.6	0.6	78	69	8.6	3.00	6
E*Trade	1.250	0.782	140.6	21.0	47	91	−72.4	8.37	8
Verticalnet	0.941	0.635	57.3	7.0	52	66	−96.3	14.73	3
Ameritrade	0.791	0.490	45.5	7.9	50	72	−68.7	8.70	0
Yahoo!	0.730	0.401	63.8	14.2	41	60	−93.2	9.52	28
Healtheon	0.593	0.446	40.5	6.6	58	83	−79.0	10.49	1
Etoys	0.610	0.435	32.4	8.9	58	79	−99.3	13.41	14
Lycos	0.609	0.468	13.5	5.8	49	96	−49.4	9.34	19
Priceline	0.565	0.337	48.5	12.9	46	65	−97.5	12.40	21
Ticketmaster	0.377	0.466	4.7	0.7	80	115	−79.6	10.01	4
Amazon	0.401	0.247	103.8	16.8	66	76	−81.1	10.78	37
Go2net	0.366	0.409	4.8	0.9	49	76	−63.1	10.24	3
CNet	0.370	0.339	12.3	3.7	49	73	−74.8	10.24	10
Webvan Group	0.319	0.303	20.2	3.1	84	96	−97.2	15.12	8
E-Bay	0.245	0.155	28.3	3.9	49	62	−74.5	10.81	63
MP3.com	0.142	0.056	14.4	5.4	68	128	−89.8	15.80	22

[a]Bullishness refers to the unweighted Naive Bayes classification.
[b]Activity is measured in thousands of messages.
[c]Intensity is measured as the average number of words per message.
[d]Return is the change in price between the first and the last trading day of the year. Lycos and Go2net stopped trading in late October.
[e]Average of the daily volatility measure which has been calculated as 1,000 times the standard deviation of the MA(1)-demeaned log price changes between 15-minute intervals.
[f]RB and YF indicate Raging Bull and Yahoo! Finance.

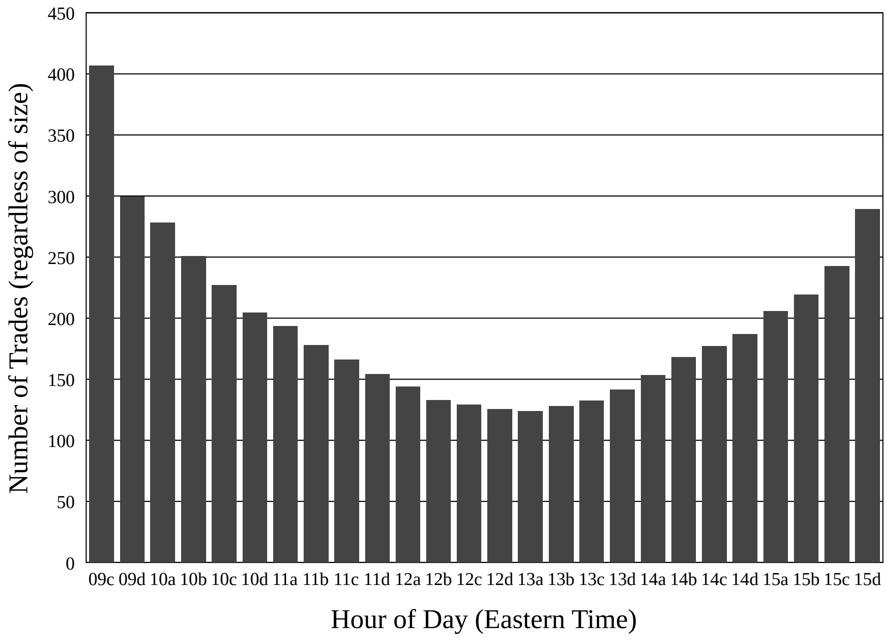

Figure 4. Hourly distribution of stock trades. Number of trades of all 45 companies in DIA and XLK combined. The letters a–d on the right side of an hour indicate fifteen minute intervals within each hour.

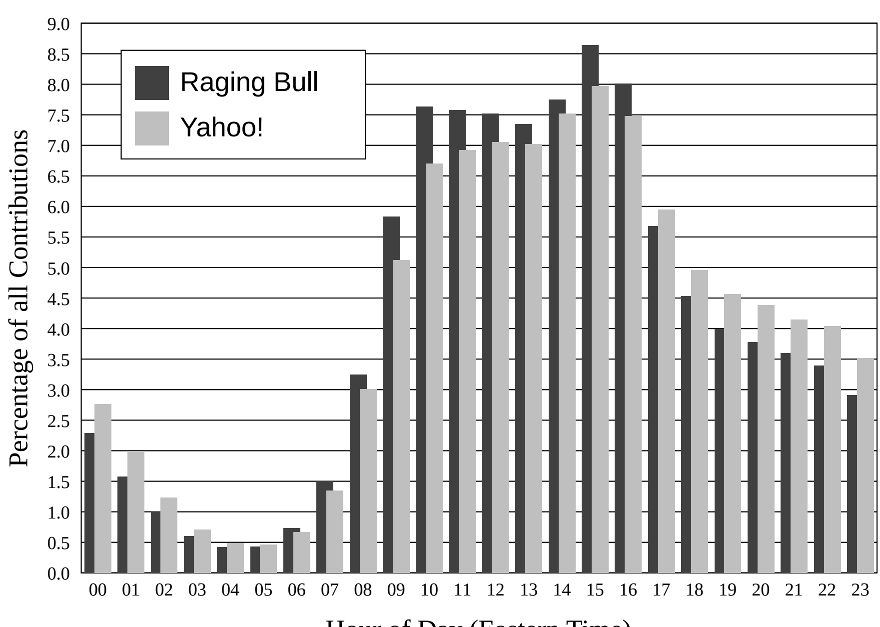

Figure 5. Hourly distribution of message board activity. Percentage of postings of 45 companies in DIA and XLK combined.

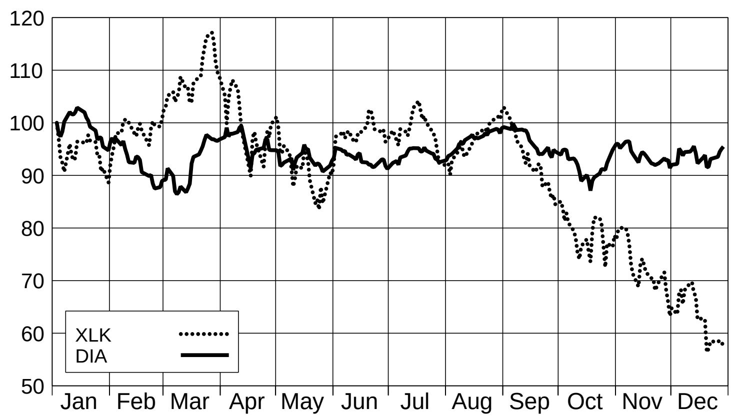

Figure 6. Stock index performance. Daily indices for DIA and XLK were computed by using the prices of the corresponding tracking funds. The indices were set to 100.0 for the first trading day of 2000. The composition of the XLK index was adjusted when stocks ceased trading: Go2net on October 13, and Lycos on October 30.

Figure 6 depicts the market performance of our versions of the DIA and XLK indices.[14] The DIA declined by about 5%, while the XLK declined by more than 40%. The decline in the XLK reflects what is often described as the end of the Internet bubble. Two of the firms in the XLK ceased trading during October due to mergers.

Figure 7 depicts movement of the bullishness signals for the DIA and XLK indices. Weekly average values are shown because the daily, and higher frequency, values are much more volatile. The bullishness signals for the XLK held up remarkably strongly. Gradual decline was observed during the year.

A popular aphorism during the so-called Internet bubble was "buy on the dips." There were many reports in the business press that at least some investors believed that following a "buy on the dips" strategy would be key to building long-term wealth. Of course, this saying does not specify a particular definition of a dip, nor does it specify how much to buy. Nonetheless it may help explain the degree to which bullishness held up in the face of declining markets.

Comparison between Figures 6 and 7 is interesting. One might imagine that stock price declines would be matched by declines in bullishness. There were two very sharp declines in XLK, one from late March to late May, and one from early September through the end of year. There is no simple tight connection between the declining stock prices and the associated bullishness indices that is apparent in these figures. There does seem to have been a decline in bullishness during the fall relative to the earlier part of the year.

[14] The figure depicts the prices of the exchange-trade tracking funds, with prices standardized at 100 at the beginning of 2000. The XLK fund imposes an upper limit of 10% on the weight of a particular stock. If market-cap weights had been used, the decline of the Internet commerce stocks would have been more dramatic due to the steep declines of the two largest stocks.

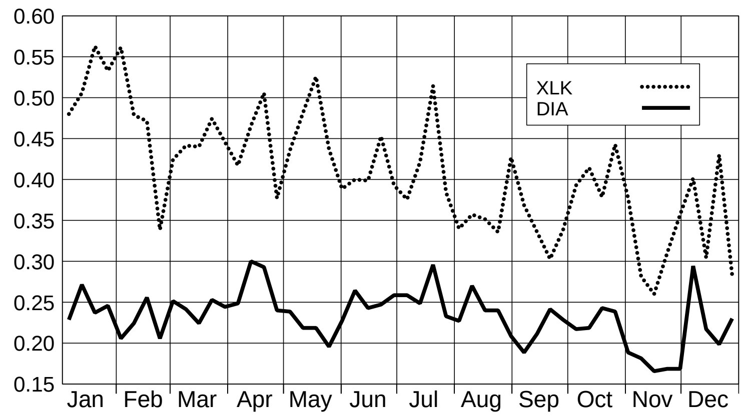

Figure 7. Message board bullishness indicator. The (unweighted) bullishness signal is based on the Yahoo! Finance message board only. Due to the high intraweek volatility of the bullishness signal, the above chart depicts weekly averages corresponding to the DIA and XLK indices. Unlike Figure 6, fixed marketcap weights were used in constructing the indices.

E. Descriptive Statistics

Table II reports a number of descriptive statistics[15] for the companies in our sample. Comparing Yahoo! Finance to the Raging Bull results in the technical appendix, we find similar results. There are more messages posted for all firms, and the messages are often more bullish. The average message is longer on Raging Bull. In comparison to the XLK firms, the DIA firms have lower losses, lower volatility, lower activity levels on the stock message boards, and less bullishness expressed on the stock message boards.

The coverage of our sample of firms in the WSJ is not all that highly correlated with message posting. The XLK firms generate a great deal of message posting, but a much lower level of coverage by the WSJ.

The average price[16] at which our sample of stocks traded during the period under study was \$58.25. The highest price observed in the data was \$496.03 and the lowest price observed was \$0.17. This range is unusually large for firms that are included in major stock indices.

While stock prices may be set efficiently, there is no economic force that would cause messages posted on Internet stock message boards to be efficient. Accordingly, we did not expect the messages to provide unbiased forecasts, and they are not unbiased. While some day traders may be active on the message boards, it seems unlikely that market makers or most institutional investors

[15] In order to compute elasticities and to control for scaling, a number of the key variables are log transformed. In particular, variables related to trading volume and to the number of messages posted are calculated as $\ln(1 + x)$ in order to avoid taking the log of zero when x is zero.

[16] In order to control for bid–ask bounce, Jones, Kaul, and Lipson (1994) recommend the use of the midpoint of the bid–ask spread rather than transactions prices. The two approaches lead to very similar results, and report results based on their recommendation.

do much message posting. People who hold either a long position or a short position are likely to be particularly interested in a given stock. People who hold a zero position may be less likely to pay attention to a particular message board. Current institutions make it much easier for a small trader to hold a long position than a short position. Thus we expected to find a bullish tone on average. This is what we find.

III. Relating Message Boards and Stock Returns

The central relations between the message boards and the stock market are provided by Tables III–V . These tables present correlations, contemporaneous regressions and time sequencing regressions. We first explain the structuring of these tables, and then organize the discussion of the results by substantive topic rather than by table.

Table III provides simple pairwise correlations that are significantly different from zero. The correlations reported at the top of Table III are among the financial measures. Most of correlations in Table III reproduce results that are already well known from studies of earlier time periods. The correlations between the stock market features and the message board features are large. In many cases these correlations are larger than the heavily studied correlations between different stock market features. For instance, the heavily studied correlation between trading volume and volatility is 0.063 in our sample. The correlation between trading volume and the number of messages is 0.322, while the correlation between the number of messages and volatility is 0.132. In other words, the magnitudes of the relationships between stock market attributes and message board measures are not trivial.

Table IV provides contemporaneous regressions using panel regressions with fixed effects for each company. The market is used as an added control factor. Time periods are 15 minutes. Each of the financial variables is treated as a dependent variable, and is regressed on a set of company dummies and three variables from the message boards: The log number of messages, the bullishness index B^*, and the agreement index A.[17] We also employ a stock market index in the regressions to control for overall market-wide effects.[18]

[17] The bullishness and agreement indices are contemporaneous with the financial variables in Table IV. An alternative approach is to use the one-hour, four-hour, and one-day lagged aggregates of these indices. When this approach is taken the magnitude and significance of the coefficients on bullishness and agreement increase. Perhaps more interestingly, the observation that the coefficient on bullishness declines with increasing trade size becomes more pronounced. This would reinforce the hypothesis that day traders play a particularly significant role on the message boards.

[18] In the regressions that explain stock returns, both theory and evidence show that it is the difference of log SPY (i.e., return) that belongs in the regression, and so we do that. In the other regressions it is less clear whether to include the log SPY or the difference of the log SPY. Prior literature does not provide clear guidance on this issue. Empirically log SPY performs markedly better in these regressions, and so we report these results. The results concerning the effects of the message boards do not depend on which approach is taken. For our purposes this is a side issue. However, it is actually an issue that might merit further study in its own right. It is not obvious to us why the market index performs better in log form rather than in the difference of log form.

Table III

Pairwise Correlations for TAQ and Yahoo! Finance

Only correlations that are significantly different from zero at the 99% confidence level are reported. Missing values are deleted on a case-by-case basis rather than for the entire data set. The time dummies for the first and last hour of the trading day are reported. Midday time dummies tended to be insignificant and economically small. The time period length is 15 minutes. Message board variables are aggregated over one hour (1 h lag), four hours (4 h lag), and 24 hours (1 d lag) prior to the current 15-minute interval, which is referred to as "no lag." The seven financial dependent variables are: the log difference in the bid–ask midpoint from the end of the previous 15-minute interval to the current 15-minute interval (return); the 15-minute price volatility; the number of small (< \$100k), medium (\$100k–\$1m) and large (>\$1m) trades; the number of traded shares (volume); and the daily average of the bid–ask spread.

	Return	Volatility	Small	Medium	Large	Volume	Spread
Volatility	−0.016						
Small trades	0.009	0.220					
Medium trades	0.023	−0.182	0.544				
Large trades	0.016	−0.093	0.421	0.677			
Volume	0.016	0.064	0.750	0.742	0.654		
Spread		−0.035	−0.086	0.064	0.043	−0.031	
09:30 Dummy	0.019	0.119	0.121	0.060	0.104	0.140	0.037
09:45 Dummy	−0.011	0.072	0.061	0.065	0.049	0.075	0.011
10:00 Dummy	−0.006	0.046	0.052	0.058	0.043	0.067	0.007
10:15 Dummy	−0.007	0.026	0.038	0.041	0.027	0.047	
15:00 Dummy			0.012	0.009		0.007	
15:15 Dummy			0.023	0.021	0.008	0.023	
15:30 Dummy		0.016	0.041	0.040	0.019	0.047	
15:45 Dummy		0.032	0.076	0.079	0.059	0.096	
Messages, no lag	−0.009	0.133	0.485	0.147	0.175	0.322	−0.069
Messages, 1h lag	−0.012	0.123	0.502	0.143	0.173	0.326	−0.072
Messages, 4h lag	−0.009	0.099	0.475	0.126	0.156	0.300	−0.072
Messages, 1d lag	−0.009	0.109	0.520	0.180	0.212	0.352	−0.075
Words, no lag	−0.008	0.077	0.347	0.117	0.146	0.253	−0.055
Words, 1h lag	−0.011	0.091	0.422	0.136	0.170	0.302	−0.068
Words, 4h lag	−0.009	0.078	0.432	0.135	0.170	0.304	−0.072
Words, 1d lag	−0.008	0.098	0.509	0.191	0.225	0.373	−0.078
Bullishness, no lag		0.124	0.334	0.025	0.060	0.192	−0.052
Bullishness, 1h lag	−0.007	0.147	0.416	0.029	0.083	0.238	−0.068
Bullishness, 4h lag	−0.008	0.133	0.419	0.028	0.090	0.239	−0.072
Bullishness, 1d lag	−0.005	0.148	0.406	0.037	0.104	0.253	−0.063
Agreement, no lag		0.119	0.327	0.035	0.073	0.191	−0.053
Agreement, 1h lag		0.116	0.326	0.043	0.092	0.197	−0.052
Agreement, 4h lag		0.066	0.201	0.034	0.071	0.136	−0.034
Agreement, 1d lag			−0.024	−0.022			

Table V provides time sequencing tests (see Hamilton (1994)) to check whether stock message board changes precede financial market changes. This table controls for a market factor and uses a dummy variable for the start of the trading week. The results are based on panel regressions with fixed effects for each company. In the technical appendix we also include variables to reflect the presence of news stories in the WSJ.

Table IV

Contemporaneous Regressions

The units of observation are the 15-minute intervals between 09:30 and 16:00 on trading days. All regressions use company fixed effects. A coefficient that is significant at the 95% level is indicated with superscript a, while superscript b and superscript c denote significance at the 99% level and 99.9% level, respectively. Absolute t-statistics are shown in parentheses. The regressors were obtained from the message boards: the log transformation $\ln(1 + M_t)$ of the number of messages; the bullishness measure B_t^* and the agreement index $A_t \in [0, 1]$. The seven financial dependent variables are the log difference in the bid–ask midpoint from the end of the previous 15-minute interval to the current 15-minute interval (return); the percentage ratio of 15-minute price volatility relative to the interval's average price; the log number of small (<\$100k), medium (\$100k–\$1m) and large (> \$1m) trades; the log number of traded shares (volume); and the daily average of the bid–ask spread. The log number of trades and volume are calculated as $\ln(1 + x)$. Market denotes the log price of the S & P 500 tracking fund (SPY), except in the case of the return regressions, where it denotes the return (difference of the log price) of the SPY.

	Log of Messages		Bullishness Index		Agreement Index		Market		R^2
Return	−0.331	(1.382)	1.747	(3.208)	−0.240	(0.455)	0.716^c	(120.7)	0.049
Volatililty	0.041^c	(35.7)	0.033^c	(12.74)	-0.029^c	(11.41)	-1.178^c	(81.85)	0.538
Log small trades	0.225^c	(102.1)	0.181^c	(36)	-0.123^c	(25.3)	-1.541^c	(55.88)	(0.984)
Log medium trades	0.119^c	(43.53)	0.161^c	(25.82)	-0.096^c	(15.84)	-0.464^c	(13.55)	(0.931)
Log large trades	0.082^c	(37.29)	0.052^c	(10.39)	-0.021^c	(4.382)	-0.222^c	(8.073)	(0.642)
Log trading volume	0.259^c	(82.37)	0.170^c	(23.81)	-0.109^c	(15.72)	-2.417^c	(61.55)	(0.995)
Spread	0.001	(0.766)	0.009^b	(2.861)	−0.004	(1.369)	-0.047^b	(2.763)	0.245

We were concerned about the extent to which our results might have been affected by sparseness of message postings. To address this concern, we repeated the daily level tests for the sample of 10 firms with the greatest posting levels.[19] While some minor changes are observed, the basic patterns are very similar. To save space we do not report the results.

Stock returns are known to be difficult to predict. Can the message boards predict returns? Table III shows that there is a significant but negative contemporaneous correlation between number of messages and stock returns. As shown in the technical appendix, this is also found with respect to the Raging Bull message boards. The same effect is present in the contemporaneous regressions, although in Table IV the result is not statistically significant.

For trading purposes, the ability to predict subsequent changes is more important than contemporaneous correlations. The time-sequencing analysis in

[19] The top 10 firms are listed in order followed by ticker, number of messages posted (Yahoo Finance + Raging Bull), and number of words posted (Yahoo Finance + Raging Bull): (1). Microsoft 197530 12301312; (2). E*Trade 161649 8576062; (3). Amazon 120574 8154611; (4). Intel 94347 5356805; (5). Philip Morris 84198 6491652; (6). Yahoo 78008 3465359; (7). AT&T 76769 4394984; (8). Verticalnet 64293 3449676; (9). Priceline 61397 3052542; and (10). Boeing 56279 4593476.

Table V

Time Sequencing Tests

The results in this table are based on messages obtained from the Yahoo! Finance message boards, aggregated to daily frequency. We estimate each equation as a panel with company fixed effects. The regressors X_i and Y_i are subscripted by their lags. The notation SPY is a variable with the log of the price of the Standard & Poors Depositary Receipt S&P 500 Tracking Fund, except in the return regressions where this variable is the time-differenced log of the price. The notation NWK is an indicator variable for a day being the first trading day after a weekend or holiday. A coefficient that is significant at the 99% level is indicated with superscript a, while superscript b and superscript c denote significance at a 99.9% level and a 99.99% level, respectively. The four X variables were obtained from the message boards. The messages and words variables were transformed into logarithms, while bullishness and agreement are the B_t^* and A_t measures defined in the text. The seven financial Y variables are the log differences in daily closing price (return); the percentage ratio of daily price volatility relative to the day's average price; the log number of small (<\$100k), medium (\$100k–\$1m) and large (>\$1m) trades; the log number of traded shares (volume); and the daily average of the bid–ask spread.

		$Y = f(X_{-1}, X_{-2}, \text{NWK}, \text{Market})$					$X = f(Y_{-1}, Y_{-2}, \text{NWK}, \text{Market})$				
X	Y	X_{-1}	X_{-2}	NWK	Market	χ^2	Y_{-1}	Y_{-2}	NWK	Market	χ^2
Messages	Return	−0.002[a]	0.002[a]	−0.002	0.096[c]	11.2[a]	−0.083	0.106	−0.525[c]	−0.332	1.34
Messages	Volatility	0.015[c]	−0.010[b]	−0.013[a]	−0.557[c]	22.0[c]	−0.050	0.055	−0.527[c]	−0.328	4.48
Messages	Small	0.074[c]	−0.027[c]	−0.043[c]	−0.507[c]	200.[c]	−0.060[a]	0.094[c]	−0.535[c]	−0.293	25.5[c]
Messages	Medium	0.049[c]	−0.051[c]	−0.100[c]	0.209	55.4[c]	−0.029	0.011	−0.531[c]	−0.318	7.59
Messages	Large	0.100[c]	−0.067[c]	−0.206[c]	0.123	96.6[c]	−0.021	0.010	−0.529[c]	−0.409	5.50
Messages	Volume	0.111[c]	−0.029[c]	−0.156[c]	−0.987[c]	288.[c]	−0.030	0.056[b]	−0.532[c]	−0.276	14.9[b]
Messages	Spread	0.002	−0.002	−0.000	−0.042	2.05	−0.029	−0.034	−0.525[c]	−0.330	0.86
Words	Return	−0.001	0.001[a]	−0.002	0.096[c]	7.44	−0.275	0.244	−0.597[c]	−0.753[b]	3.89
Words	Volatility	0.005	−0.003	−0.008	−0.558[c]	5.62	−0.005	0.064	−0.598[c]	−0.688[a]	2.76
Words	Small	0.025[c]	−0.006	−0.018	−0.489[c]	64.4[c]	0.012	0.123[c]	−0.595[c]	−0.601[a]	34.8[c]
Words	Medium	0.017[b]	−0.018[c]	−0.083[c]	0.204	18.9[c]	−0.016	0.017	−0.601[c]	−0.762[b]	0.89
Words	Large	0.036[c]	−0.022[b]	−0.176[c]	−0.075	33.5[c]	0.006	0.013	−0.598[c]	−0.736	2.08
Words	Volume	0.043[c]	−0.005	−0.125[c]	−0.936[c]	122.[c]	0.040	0.082[b]	−0.594[c]	−0.498	35.4[c]
Words	Spread	0.001	−0.000	0.001	−0.042	0.26	−0.009	−0.010	−0.598[c]	−0.761[b]	0.03
Bullishness	Return	−0.002	−0.003	−0.003	0.098[c]	2.83	−0.036	0.018	−0.008	0.204[c]	1.15
Bullishness	Volatility	0.038[b]	−0.026[a]	−0.002	−0.565[c]	17.4[b]	0.018	0.008	−0.008	0.238[c]	9.51[a]
Bullishness	Small	0.136[c]	−0.064[c]	0.006	−0.534[c]	78.9[c]	0.053[c]	0.003	−0.005	0.299[c]	137.[c]
Bullishness	Medium	0.117[c]	−0.144[c]	−0.062[c]	0.209	49.2[c]	0.016[c]	−0.001	−0.007	0.198[c]	35.5[c]

(continued)

Table V—*Continued*

X	Y	$Y = f(X_{-1}, X_{-2}, \text{NWK}, \text{Market})$					$X = f(Y_{-1}, Y_{-2}, \text{NWK}, \text{Market})$				
		X_{-1}	X_{-2}	NWK	Market	χ^2	Y_{-1}	Y_{-2}	NWK	Market	χ^2
Bullishness	Large	0.209[c]	−0.076	−0.139[c]	−0.236	46.6[c]	0.007	0.005	−0.009	0.044	17.2[b]
Bullishness	Volume	0.185[c]	−0.045	−0.086[c]	−1.013[c]	83.4[c]	0.030[c]	0.005	−0.008	0.296[c]	73.3[c]
Bullishness	Spread	0.016[a]	0.000	0.002	−0.048	11.2[a]	0.024	0.024	−0.009	0.202[c]	4.75
Agreement	Return	−0.002	−0.008	−0.003	0.098[c]	4.93	−0.010	0.012	0.003	0.162[c]	0.39
Agreement	Volatility	0.026	−0.024	−0.003	−0.560[c]	4.51	0.013	0.001	0.003	0.180[c]	7.63
Agreement	Small	0.091[b]	−0.054	0.004	−0.506[c]	16.7[b]	0.026[c]	0.003	0.005	0.210[c]	104.[c]
Agreement	Medium	0.106[a]	−0.138[b]	−0.064[c]	0.204	19.2[c]	0.010[c]	0.002	0.004	0.157[c]	48.1[c]
Agreement	Large	0.226[c]	−0.050	−0.141[c]	−0.231	21.9[c]	0.004	0.004	0.003	0.132	22.3[c]
Agreement	Volume	0.142[c]	−0.057	−0.089[c]	−0.953[c]	21.5[c]	0.013[c]	0.004	0.004	0.206[c]	.0[c]
Agreement	Spread	0.028[b]	−0.001	0.002	−0.049	13.3[a]	0.018	0.033[a]	0.003	0.161[c]	14.1[b]

Table V shows a significant negative relationship between the number of messages on day t, and stock returns on day $(t + 1)$. On the subsequent day $(t + 2)$ the effect reverses itself. Corresponding analysis in Table V of our technical appendix for 15-minute intervals provides further support for the presence of negative predictability.

The next-day reversal looks much like a bid–ask bounce. In an effort to control for bid–ask bounce, we follow Jones et al. (1994) in using the midpoint of the bid–ask spread as the underlying price from which the returns are calculated. Either their procedure provides an incomplete removal of the bid–ask bounce, or else there is some other effect at work. However, since the magnitude is economically very small, we do not try to disentangle these two interpretations.

Statistical significance does not imply economic importance. The returns predictability is a statistically significant effect, but it is very small in magnitude. The estimated effect is -0.002 for the past day and $+0.002$ for the second-last day. Thus the two effects cancel out. The partial effect of -0.002 implies that a doubling of the number of messages (a 100% increase) leads to a mere -0.2% decrease in the stock price. The size of the effect is larger for 15-minute intervals, where our estimate indicates an elasticity of -0.004 for the most recent interval. However, even here such tiny price changes would be difficult to take advantage of because potential gains would likely be offset even by small transaction costs.

Is bullishness associated with returns? In the contemporaneous regressions reported in Table IV, greater bullishness is positively and significantly associated with returns. In the time sequencing tests reported in Table V, bullishness is not statistically significant.

IV. Volatility

If the people posting messages are the real-world counterpart of "noise traders," then their actions might induce market volatility. Furthermore, we know that volatility and volume are correlated. We also know that message posting and trading volume are correlated. For these reasons we next consider whether the messages help forecast volatility. There are many approaches to market volatility. We have used methods based on the currently popular "realized volatility" approach as our main case. However, we have also put effort into GARCH class models in order to ensure robustness.

French, Schwert, and Stambaugh (1987) and Schwert (1990) use actual within-period variation of returns to measure the volatility of a particular period. Theoretical justification of this approach is provided by Andersen et al. (2001), who use the term "realized volatility" to describe the approach. Realized volatility is attractive because it makes volatility essentially observable, so it can then be used in standard econometric procedures. Andersen, Bollerslev, Diebold and Ebens (2001) have shown the value of this approach in a study of the firms in the Dow Jones Industrial Average. For this reason we use a realized volatility approach to investigate the impact of the message boards.

The GARCH class of volatility models has been extremely influential over many years. In a recent comparative volatility study, Hansen and Lunde (2001) include 330 GARCH class methods. As shown by Hansen and Lunde, it turns out to be very hard to outperform a traditional GARCH (1,1). The main defect of the GARCH (1,1) for stock volatility is that this model does not allow for asymmetric responses to shocks. This asymmetry is known as the "leverage effect." In a recent analysis of the Dow Jones Industrial Average, Engle and Patton (2001) report results for the GARCH (1,1) specification as well as a GJR (1,1,1) that is based on Glosten et al. (1993). Thus we considered GARCH, EGARCH, and GJR models. In order to save space these results are reported in our technical appendix.

It is well known that trading volume helps to predict volatility (see Jones, Kaul, and Lipson (1994)), and thus we include trading volume as an added factor. Glosten et al. (1993) and Engle and Patton (2001) consider the use of the Treasury bill rate as an added regressor. Glosten et al. argue that it proxies for inflation risks. Given that inflation was relatively tame during 2000, it is perhaps not surprising that when we tried adding a Treasury bill rate, it was not significant. Thus we do not separately report these results.

Our implementation of the realized volatility method follows Andersen, Bollerslev, Diebold, and Ebens (2001) very closely. The main difference is that for the daily data, we employ a 15-minute intraday period. First, we fit an MA(1) model for each stock using the full year 2000 of data as in Andersen et al. (2001). We remove the MA(1) component for each stock giving us a filtered series of 15-minute returns. Let the filtered series be denoted by $r_{i,t,d}$ for stock i on day t and intra-day period d. Let $\mathcal{D}(t)$ denote the set of intra-day periods in day t, and let $|\mathcal{D}(t)|$ denote the size of the set $\mathcal{D}(t)$. Company i's standard deviation of intra-day returns on day t is thus given by

$$v_{i,t} = \left[\frac{1}{|\mathcal{D}(t)|} \sum_{d \in \mathcal{D}(t)} r_{i,t,d}^2 \right]^{0.5}. \tag{9}$$

We use the log of the standard deviation $\ln(v_{i,t})$ as our measure of volatility.

As discussed by Engle and Patton (2001) and Andersen et al. (2001), there is significant evidence that volatility is fairly long-lived. It is often thought that volatility declines hyperbolically rather than exponentially.[20] A convenient way to model this long-run dependence is through the use of fractional integration. A long memory series will be integrated as $I(d)$ with $0 < d < 0.5$. Geweke and Porter-Hudak (1983) developed a log-periodogram method of estimating d. We use this method to estimate the value of d for each series. For our stocks the observed value of d is always greater than zero and usually less than 0.5. Using estimates of the periodogram $\psi_{i,k}$ for company i at frequency points

[20] This fact has motivated the development of FI-GARCH (fractionally integrated GARCH) and H-GARCH (hyperbolic GARCH). We do not try these approaches since we have only one year of data.

$k = 0, 1, \ldots, \lfloor T^{0.6} \rfloor$, we finally estimate the fractional integration parameter from the random-effects panel regression on the estimating equation

$$\ln(\psi_{i,k}) = \alpha - d \ln\left[4 \sin\left(\frac{k\pi}{T}\right)\right] + \mu_i + u_{i,k}. \tag{10}$$

The estimate we obtain from this panel log-periodogram regression is $d = 0.297$ with a standard error of 0.021. This value is quite close to the median value of 0.349 reported by Andersen et al. (2001).

Our central interest is in understanding the extent to which the message boards have an impact on the realized volatility. To do this we estimate a "news impact function" that follows equation (5) from Andersen et al. (2001). Because of our focus we introduce trading volume and the stock message board variables. Accordingly we estimate

$$\begin{aligned}(1 - L)^d \ln(v_{i,t}) = \beta_i &+ \beta_{v,i} \ln(v_{i,t-1}) + \beta_{vr,i} \ln(v_{i,t-1}) I(r_{i,t-1} < 0) + \beta_{A,i} A_{i,t} \\ &+ \beta_{M,i} \ln(1 + M_{i,t-1}) + \beta_{N,i} \ln(N_{t-1}) + u_{i,t}\end{aligned} \tag{11}$$

with i.i.d. error term $u_{i,t}$. In addition to estimating this equation on a firm-by-firm basis, we also estimate it as a panel with company-specific random effects μ_i:

$$\begin{aligned}(1 - L)^d \ln(v_{i,t}) = \beta &+ \beta_v \ln(v_{i,t-1}) + \beta_{vr} \ln(v_{i,t-1}) I(r_{i,t-1} < 0) + \beta_A A_{i,t} \\ &+ \beta_M \ln(1 + M_{i,t-1}) + \beta_N \ln(N_{t-1}) + \mu_i + u_{i,t}.\end{aligned} \tag{12}$$

When this kind of a model is estimated on a firm-by-firm basis, cross-sectional effects are ruled out by assumption. However, we anticipated that cross-sectional differences might be important. Unlike fixed-effects and firm-by-firm regressions, the random-effects regression permits parameter estimates to be influenced by both cross-sectional and time-series variation in the data.

In order to establish the empirical context for the news impact functions, we first discuss the results from the simpler procedures in Tables III–V. While Andersen et al. (2001) only consider realized volatility at daily frequency, in Table III we are interested in looking at 15-minute frequency. Therefore, in this part we define volatility as simply the unfiltered within-period standard deviation of actual trade prices.

Table III shows that message posting is strongly positively correlated with volatility. This holds up in the contemporaneous regressions (Table IV). More messages are posted during volatile periods and the messages are somewhat more bullish. Greater contemporaneous disagreement is found during volatile periods.

Table V shows that we can predict raw volatility using the message board data. More messages today imply significantly greater market volatility tomorrow. The more bullish today's messages, the greater tomorrow's market

Table VI

By-Company Volatility Regressions and Estimates Distribution

The dependent variable in all regressions is the fractionally differenced log of the volatility measure $(1-L)^d \ln(v_{it})$ for company i at time t, which in turn has been calculated as the standard deviation of MA(1)-demeaned intra-day returns. The fractional integration parameter d has been estimated (using the full panel with random company effects) as 0.3097. The notation $\ln(v_{i,t-1})$ is the lagged log of the (untransformed) volatility measure, and $I(r_{i,t-1} < 0)$ is an indicator variable for yesterday's return to be negative. The notation $\ln(N_{i,t-1})$ is the log-transformed total number of trades on the previous day, and $\ln(M_{i,t-1})$ is the log-transformed total number of messages posted during the 24 hours prior to market opening at 09:30. The agreement index $A_{i,t-1}$ refers to the same time period definition. Estimates were obtained from 45 individual OLS regressions. The last two lines show averages for the group of 30 DIA and 15 XLK companies, respectively.

	$\ln(v_{t-1})$		$\ln(v_{i,t-1}) \times I(r_{i,t-1}<0)$		Messages $\ln(1+M_{i,t-1})$		Trades $\ln(N_{i,t-1})$		Agreement $A_{i,t-1}$	
Stock	Est.	*t*-stat	Est.	*t*-stat	Est.	*t*-stat	Est.	*t*-stat	Est.	*t*-stat
Minimum	−0.264	−3.840	−0.010	−2.490	−0.062	−1.612	−0.281	−2.929	−0.213	−2.632
0.10	−0.207	−2.803	−0.004	−0.890	−0.026	−0.743	−0.051	−0.887	−0.156	−1.669
0.25	−0.149	−2.071	0.001	0.263	−0.003	−0.064	0.086	1.035	−0.064	−0.972
Median	−0.114	−1.637	0.006	1.319	0.026	0.834	0.131	1.559	−0.005	−0.061
0.75	−0.035	−0.400	0.011	2.392	0.071	1.666	0.221	2.058	0.039	0.568
0.90	0.011	0.169	0.014	3.088	0.121	2.990	0.359	2.870	0.099	1.098
Maximum	0.132	2.126	0.027	5.150	0.251	5.461	0.431	3.450	0.322	3.339
Mean	−0.097	−1.352	0.006	1.200	0.040	0.975	0.141	1.294	−0.009	−0.097
Std.Dev.	0.087	1.248	0.007	1.604	0.061	1.443	0.149	1.395	0.095	1.122
Mean (DIA)	−0.126	−1.775	0.003	0.827	0.037	0.891	0.175	1.490	−0.009	−0.183
Mean (XLK)	−0.038	−0.506	0.011	1.946	0.045	1.143	0.073	0.903	−0.008	0.074

volatility. Neither the number of words nor the extent of disagreement today forecast tomorrow's volatility.

Table V also allows a comparison of the strength of effects in each direction. In data measured both at daily frequency, and in data measured at 15-minute frequency, message posting activity has a more significant effect on market volatility than market volatility has on message posting. Thus while there is a flow in each direction, the more significant flow is from the message boards.

Table VI provides the news impact function results from individual firm estimates corresponding to (11).[21] We use Huber/White standard errors, which are quite conservative, and we add the log number of trades as an explanatory factor.[22] The estimates for β_v and β_{vr} are fairly close although not identical to Andersen et al. (2001). This is not surprising since our sample of firms differs by the inclusion of the 15 firms in the XLK index. The time periods are also different. We find less evidence of the asymmetry known as the "leverage effect," it is more important for some firms than for others. Both the number

[21] These results can also be directly compared with Table V in Andersen et al. (2001).

[22] As in Chan and Fong (2000) and Jones, Kaul, and Lipson (1994), it is well known that the number of trades helps forecast volatility.

of stock messages and the number of trades exhibit significant cross-sectional differences and are not quite significant at the mean. Agreement among the messages appears to be insignificant.

For a number of the well-known high tech firms, including Amazon, IBM, MP3.com, and Microsoft, the number of messages posted has a significant positive effect on volatility. E-Bay is the only firm for which agreement plays a significant negative effect on volatility.

Table VII estimates a random-effects pooled regression version of Table VI. We consider all-firm regressions as well as separate regressions for the DIA and XLK firms. The importance of the cross section differences is clear. For the DIA firms, $\beta_{v,i}$ is on average negative. But for the XLK firms the effect is positive and significant. When we pool the larger number of DIA firms with the smaller number of XLK firms, there is a negative aggregate effect. There is stronger asymmetry observed among the XLK firms than among the DIA firms.

Tables VI and VII show that the connection between the number of trades and market volatility is more complex than is sometimes recognized. The number of trades is a significant factor for the DIA firms, but it is not significant for the XLK firms. The number of messages is a predictive factor of volatility for both groups of firms. Agreement among the posted messages is not significant.

Message posting is clearly significant in the cross section; it has a smaller magnitude for the DIA firms than it does for the XLK firms. These findings are consistent with the descriptive statistics in Table II. Recall that the XLK firms are more volatile than the DIA firms and recall that they also typically have higher message posting levels. Thus the finding that message posting has a greater impact on volatility for the XLK firms is perhaps not unexpected.

V. Trading Volume

Two aspects of trading volume have received attention in the academic literature. Hirshleifer (1977), Harris and Raviv (1993), and others present a traditional hypothesis according to which disagreement causes trades to take place. Trading volume is also used as a proxy for liquidity, as in Chordia, Roll, and Subrahmanyam (2001). We consider each of these issues.

Table IV provides support for the Harris and Raviv (1993) hypothesis that disagreement produces trading. These regressions control for general movements in the overall market and explain all trades as well as trades of different sizes. We find that there are negative coefficients on the agreement index. This effect is most marked for the small trades, but it remains significant for larger trades, although the magnitudes are economically smaller. The explanatory power of the volume regressions is quite high.

Table V shows that the message boards can be helpful in predicting trading volume. In data measured at daily frequency, the effect from message posting activity to trading volume is a more significant effect than is the reverse direction. As shown in the technical appendix, in data measured at 15-minute frequency, the effect from trading volume to the number of messages posted is the more significant direction.

Table VII

Volatility Panel Regressions

The dependent variable in all regressions is the fractionally differenced log of the volatility measure $(1-L)^d \ln(v_{it})$ of company i at time t, which in turn has been calculated as the standard deviation of MA(1)-demeaned intra-day returns. The fractional integration parameter d has been estimated (using the full panel with random company effects) as 0.3097. The notation $\ln(v_{i,t-1})$ is the lagged log of the (untransformed) volatility measure, and $I(r_{i,t-1} < 0)$ is an indicator variable for yesterday's return to be negative. The notation $\ln(N_{i,t-1})$ is the log-transformed total number of trades on the previous day, and $\ln(1+M_{i,t-1})$ is the log-transformed total number of messages posted during the 24 hours prior to market opening at 09:30. The agreement index $A_{i,t-1}$ refers to the same time period definition. The panels of 45 (all), 30 (DIA), and 15 (XLK) companies were estimated using random company effects. Pseudo-R^2 are reported in the last column. Absolute t-statistics are shown in parenthesis. Maximum-likelihood estimates were obtained from a regression with random company effects.

Panel	Intercept		$\ln(v_{t-1})$		$\ln(v_{i,t-1}) \times I(r_{i,t-1}<0)$		Messages $\ln(1+M_{i,t-1})$		Trades $\ln(N_{i,t-1})$		Agreement $A_{i,t-1}$		R^2
All	0.065	(1.16)	−0.007	(.750)	0.005[c]	(7.08)	0.039[c]	(8.82)			−0.010	(1.01)	0.118
DIA	−0.098	(1.38)	−0.039[c]	(3.48)	0.003[c]	(3.85)	0.023[c]	(4.63)			−0.011	(.990)	0.078
XLK	0.396[c]	(3.82)	0.039[a]	(2.35)	0.011[c]	(7.30)	0.025[b]	(2.65)			0.002	(.076)	0.149
All	−0.510[c]	(3.80)	−0.033[b]	(2.98)	0.005[c]	(7.67)			0.073[c]	(6.51)	−0.003	(.248)	0.052
DIA	−0.433[a]	(2.16)	−0.052[c]	(3.37)	0.003[c]	(4.15)			0.044[b]	(2.62)	−0.007	(.672)	0.074
XLK	0.486	(1.78)	0.050	(1.90)	0.011[c]	(7.59)			0.008	(.487)	0.004	(.579)	0.138
All	−0.386[b]	(2.85)	−0.033[b]	(2.92)	0.005[c]	(7.29)	0.033[c]	(6.83)	0.043[c]	(3.62)	−0.009	(.893)	0.086
DIA	−0.689	(1.80)	−0.069[c]	(3.44)	0.003[c]	(4.04)	0.022[c]	(3.34)	0.056	(1.56)	−0.011	(1.03)	0.076
XLK	0.515[c]	(4.04)	0.047[b]	(2.78)	0.011[c]	(7.21)	0.025[b]	(2.84)	−0.010	(1.20)	−0.004	(.187)	0.152

Bullishness also exhibits some predictive ability for trading volume in Table V. However, the more significant effect is from trading volume to bullishness rather than from bullishness to trading volume. This result is particularly true for the smaller-sized trades. While this can be interpreted in terms of market manipulation strategies, it is also consistent with small traders simply wishing to talk about purchases that they have just made.

The Table V direction results suggest a simple interpretation. It appears that many people post messages within minutes of conducting a trade. These results are particularly marked for the smaller-sized trades.

Next we ask whether this result survives the inclusion of the factors considered by Chordia et al. (2001) when their factors are applied to individual stocks rather than the overall market. Since we are interested in the hypothesis that small trades are more closely related to the message boards, we also decompose the trading volume into trades of different sizes. As a result, we consider a number of different dependent variables in the form of log daily changes in (a) the trading volume as defined by the number of shares; (b) the dollar trading volume; (c) the total number of trades; (d) the number of small trades less than \$10k; (e) the number of medium-size trades between \$10k and \$100k; (f) the number of large trades over \$100k.

Let p_t^o and p_t^c denote the opening and closing price (bid–ask midpoint) of a particular stock on day t. Let $\bar{p}$ denote the price of the S&P 500 spider (SPY). Further let $p_{t,d}$ denote the price at the end of a particular 15-minute interval $d \in \mathcal{D}(t)$ on day t, where consecutive trading days are joined to provide a continuous time series. We further denote by FFR, T10, and BAA the Federal Funds Rate, the 10-year constant-maturity treasury bill rate, and Moody's Baa or better corporate bond yield rate, respectively. In each case percentage figures are transformed into logarithmic form as $\ln(1 + x/100)$. Then the market-based regressors are defined as follows: Stock up Yesterday $= \max\{0, \ln(p_{t-1}^c) - \ln(p_{t-2}^c)\}$; Stock down Yesterday $= \max\{0, \ln(p_{t-2}^c) - \ln(p_{t-1}^c)\}$; Stock up Last Five Days $= \max\{0, \ln(p_{t-1}^c) - \ln(p_{t-5}^o)\}$; Stock down Last Five Days $= \max\{0, \ln(p_{t-5}^o) - \ln(p_{t-1}^c)\}$; Stock Five-day Volatility $= \sum_{s=1}^{5} \sum_{d \in \mathcal{D}(t)} |\ln(p_{t-s,d}) - \ln(p_{t-s,d-1})|$; Market up Yesterday $= \max\{0, \ln(\bar{p}_{t-1}^c) - \ln(\bar{p}_{t-2}^c)\}$; Market down Yesterday $= \max\{0, \ln(\bar{p}_{t-2}^c) - \ln(\bar{p}_{t-1}^c)\}$; Market up Last Five Days $= \max\{0, \ln(\bar{p}_{t-1}^c) - \ln(\bar{p}_{t-5}^o)\}$; Market down Last Five Days $= \max\{0, \ln(\bar{p}_{t-5}^o) - \ln(\bar{p}_{t-1}^c)\}$; Market Five-day Volatility $= \sum_{s=1}^{5} \sum_{d \in \mathcal{D}(t)} |\ln(\bar{p}_{t-s,d}) - \ln(\bar{p}_{t-s,d-1})|$; Federal Funds Rate $= \Delta\mathrm{FFR}_t$; Term Spread $= \Delta(\mathrm{T10}_t - \mathrm{FFR}_t)$; and Quality Spread $= \Delta(\mathrm{BAA}_t - \mathrm{T10}_t)$. We also employ a series of dummy variables: day-of-week dummies for Monday, Tuesday, Wednesday, and Thursday. A dummy variable indicates if a trading day is right before or after a civic holiday, except when that trading day falls on a Monday or Friday.[23]

[23] The indicator variable is one for: Tuesday, January 18 (day after Martin Luther King Day); Tuesday, February 22 (day after Presidents' Day); Thursday, April 20 (day before Good Friday); Tuesday, May 30 (day after Memorial Day); Wednesday, July 5 (day after Independence Day); Tuesday, September 5 (day after Labor Day); Wednesday, November 22 (day before Thanksgiving Day); and Tuesday, December 26 (day after Christmas Day).

In order to examine the role of the message boards, we add message board variables as the log of the number of messages posted during the last 24 hours prior to market opening, and the corresponding agreement index. We expect that greater posting volume should predict a subsequent increase in trading volume.

Results are reported in Table VIII and are mostly fairly consistent with Chordia, Roll, and Subrahmanyam (2001). The daily dummies reflect elevated trading at midweek. The interest rate plays an important role. While Chordia, Roll, and Subrahmanyam focus on market-wide movements, we find that the firms' specific movements are rather more important. We add dummy variables for the presence of relevant news stories in the WSJ. If there was an article in today's WSJ, trading is elevated. If there was an article in yesterday's WSJ, then today's trading is depressed.

The crucial issue for our purposes is whether the message board posting levels remain significant when such a large number of other factors are taken into account. It turns out that the message board posting volume remains a predictor of trading volume. The sign is positive and it is statistically significant.

When we break down the trading volume according to trade size, some interesting patterns emerge. Small trades are highest on Mondays, while large trades are not. Large trades are much more elevated on Tuesday and Wednesday than are small trades.

A number of other facts about trading volume are found in Tables V and VIII. It is well known that the first day of the trading week tends to be a bad day. This is known as the "Monday effect." One interpretation that has been occasionally offered for the Monday effect is psychological: People are more depressed when they return to work after the weekend. Lakonishok and Maberly (1990) found a reduction in trades by institutional investors and an increase in trades by small investors at the start of the week. We also find that on the first trading day of the new week there is a dramatic drop in the number of messages posted, and in the length of the messages. Many people are presumably more busy doing other things. Consistent with the psychological interpretation, the sign on bullishness is negative at the start of the new week. However, in sharp contrast to the activity effects, the change in bullishness is statistically insignificant and small in magnitude.

The possible role of agreement on trading volume is harder to predict. Table IV shows that, as expected, greater agreement in a period is associated with fewer trades during that period. What is the effect of disagreement in a period on the number of trades next period? If we suppose that it takes a while for people to get around to completing their trades ("sluggish traders hypothesis"), then we would predict that greater disagreement now would be associated with more trades tomorrow. Cao et al. (2002) suggest that when a person learns that others have received the same signal that they have received (greater agreement), this will make them more likely to trade. Since it takes some time for this agreement to be revealed, this prediction is that greater agreement on a given day will be followed by more trades on the next day. Table V provides evidence

Table VIII

Volume Regressions

The dependent variable in these panel regressions of daily company data is the log change in the variable indicated in each column head. The notation \$-Volume refers to the \$-value of the shares traded, while Small and Large indicate the number of small (<\$100k) and large (>\$1m) trades. We have excluded results for the middle group because the corresponding results were very close to the results for small trades. The regressors include company-specific and market-wide variables that measure the log change in daily closing prices (denoted by $\Delta \ln p$), as well as volatility. "Up" and "Down" indicate $\max\{0, \Delta \ln p\}$ and $\max\{0, -\Delta \ln p\}$ transformations, ensuring that these regressors are non-negative. "Yesterday" and "5 Days" indicate the length of the differencing. Volatility is computed as the sum of the absolute log price change over five days and 26 daily 15-minute intervals. Federal funds rate, term spread, and quality spread are expressed as daily differences in the $\ln(1 + q/100\%)$-transformed federal funds rate and two interest rate differentials. Term spread is the difference between the (log-transformed) 10-year constant maturity government bond rate and the federal funds rate. Quality spread is the difference between the (log-transformed) 10-year maturity Moody's Baa or better corporate bond yield and the 10-year constant maturity government bond rate. The message board variables are based on the four hour aggregates of messages posted on the Yahoo! Finance message boards prior to market opening at 09:30. The "Articles in WSJ" variables are indicators that are one if one or more articles appeared in the WSJ today, yesterday, or the day before yesterday. Coefficients were estimated using random effects on the panel of 45 companies in our data set. A coefficient that is significant at 95% level is indicated with superscript a, while superscript b and superscript c denote significance at the 99% level and 99.9% level, respectively. Absolute t-statistics are shown in parentheses.

Dependent Variable	Δ \$-Volume		Δ Small		Δ Large	
Stock up yesterday	-1.020^c	(7.12)	-1.167^c	(10.7)	-1.034^c	(4.46)
Stock down yesterday	-0.468^c	(4.20)	-0.335^c	(3.96)	-0.419^a	(2.32)
Stock last 5 days up	-0.589^c	(6.54)	-0.415^c	(5.86)	-0.522^c	(3.53)
Stock last 5 days down	-0.188^c	(3.60)	-0.104^a	2.23	-0.191^a	2.11
Stock 5 day volatility	-0.492^b	(2.87)	−0.346	(1.27)	−0.109	(.254)
Market up yesterday	0.027	(.092)	−0.244	(1.13)	−0.007	(.015)
Market down yesterday	0.049	(.530)	0.071	(1.03)	0.082	(.555)
Market last 5 days up	−0.255	(.850)	0.059	(.265)	−0.187	(.390)
Market last 5 days down	-0.284^c	(3.58)	-0.118^a	(1.97)	-0.411^b	(3.22)
Market 5 day volatility	0.190	(1.84)	0.089	(1.07)	0.060	(.352)
Federal funds rate	-41.15^c	(3.73)	-45.97^c	(5.59)	-53.74^b	(3.04)
Term spread	−19.88	(1.88)	-19.17^a	(2.43)	-36.20^a	(2.13)
Quality spread	−17.76	(1.13)	−1.569	(.134)	−24.83	(.987)
Message board volume	0.012^b	(3.05)	0.009^b	(2.91)	0.016^a	(2.55)
Agreement index	−0.015	(1.48)	−0.010	(1.36)	−0.027	(1.71)
Intercept	0.005	(.432)	−0.016	(1.32)	−0.006	(.285)
Monday	0.010	(.719)	0.111^c	(10.3)	−0.029	(1.23)
Tuesday	0.108^c	(7.88)	0.062^c	(6.02)	0.130^c	(5.92)
Wednesday	0.096^c	(7.14)	0.085^c	(8.43)	0.115^c	(5.30)
Thursday	0.026	(1.90)	0.034^c	(3.29)	0.040	(1.80)
Holiday	0.017	(.694)	0.088^c	(4.71)	−0.042	(1.04)
Articles in WSJ today	0.016	(1.89)	0.013^a	(2.00)	0.028^a	(2.01)
Articles in WSJ yesterday	-0.044^c	(5.10)	-0.029^c	(4.49)	-0.073^c	(5.29)
Articles in WSJ 2 days ago	−0.013	(1.53)	−0.011	1.71	−0.000	(.009)
σ^2 (mean squared error)	0.196^c	(74.2)	0.109^c	(73.3)	0.504^c	(73.8)
ρ (cross-section variance ratio)	0.000	(.000)	0.000	(.000)	0.000	(.001)
Observations	10,973		10,973		10,973	
Companies	45		45		45	
Pseudo-R^2	0.203		0.228		0.145	

that greater agreement on a given day is followed by more trades on the next day. This effect is empirically quite strong.

VI. Message Boards and a Traditional Source of Information

To what extent do the stock message boards merely repeat what was already known from press reports? In order to answer this question we used Lexis/Nexus to collect articles about the companies from the WSJ. Table II shows that the XLK firms receive relatively more discussion on the message boards, while the WSJ provides relatively more coverage of the DIA firms. In the technical appendix to this paper we show that the main time sequencing/predictive effects reported in Table V are robust to the inclusion of dummy variables that account for the presence to such stories. Thus the message boards are not merely repeating what is in the WSJ.

Can the message boards be used to predict the presence of news stories in the WSJ? The day of a news story we might expect to see discussion of what the news means. The day before a news story in the WSJ is a problematic day because news wires and web pages will have already posted some of the stories that will be printed next day. Two days before a news story is a cleaner day for testing predictability.

Table IX reports logit regressions in which the dependent variable is one on a day with a story and zero on a day without. In addition to the message board measures, the regressions include company and day of the week fixed effects. The coefficients in Table IX are in terms of log odds ratios. Thus the critical values of the coefficients are one. If a coefficient is greater than one that means that the odds are increased. If the coefficient is less than one that means that the odds are reduced. For example, in column 1 of the top panel, a unit increase in bullishness doubles the odds of an article in the WSJ the next day, while a doubling of the number of messages increases the odds ratio by 40% ($=100\%[1.635^{\ln 2} - 1]$).

Table IX shows that the message boards have more messages, more words, and greater bullishness on the day of a WSJ article. The same thing is true of the two days leading up to a news story. Thus unusual message board activity does predict the presence of news stories.

The business press itself has been making use of the fact that the message boards make news and rumors public very early. Colorful quotes from messages are used. The date on which rumors started to circulate about some corporate event are also reported in some stories. Direct reporting of what is said on the message boards does not seem to be sufficiently frequent to account for the results in Table IX.[24]

[24] To find current examples, go to the WSJs web page and then search for "message board." Typical examples include: "Speculation on Flooz's fate first popped up in online message boards, including dot-com rumor site F—dcompany.com and a shopping board run by Anandtech.com" (Miles (2001), online) and "Since mid-July, when rumors started to spread on Internet message boards that Fidelity was selling most of its 7% stake, the stock had dropped more than 25%." (Marcinkowski (2001), online)

Table IX

Does Posting Activity Predict News Releases in the *Wall Street Journal*

Estimates are obtained from logistic regressions with company fixed effects and day-of-week fixed effects. Based on today's posting activity, the first panel attempts to predict articles in tomorrow's WSJ, while the second panel attempts to predict articles in the WSJ the day after tomorrow. The dependent variable is a binary response variable, which is one when the WSJ has published an article on the next day (or the day thereafter in the second part of the table), and zero otherwise. In the logistic model, the log odd ratio is linear in the regressors: $\ln(z) = bx$ with $z = \text{Prob}[y]/(1 - \text{Prob}[y])$. For convenience of interpretation, estimates are presented in odds ratio form $\psi = \exp(b)$: a one-unit increase in regressor x leads to a ψ-fold increase in the odds ratio z. Where regressors are expressed in logarithmic form, an r-fold increase in the regressor translates into a $\psi^{\ln(r)}$-fold increase in the odds ratio z. Absolute t-statistics are shown in parenthesis. Significance is at the 95%, 99% and 99.9% levels indicated by superscripts a, b and c, respectively. Companies with zero or fewer than five WSJ releases in 2000 were dropped from this analysis. Also dropped were irrelevant observations when, due to a weekend or civic holiday, the WSJ is not published tomorrow (in the first panel) or the day after tomorrow (in the second panel). Message board data are from Yahoo! Finance. Log messages and log words were computed as $\ln(1 + x)$ to allow for rare cases of zero messages or words. All message board variables are aggregates over the 24-hour period of the current day.

	Does Message Posting Today Predict Articles in the WSJ Today?									
	(1)		(2)		(3)		(4)		(5)	
Messages $\ln(1+M_{it})$	1.805^{c}	(11.8)	1.767^{c}	(10.9)						
Words $\ln(1+W_{it})$					1.467^{c}	(9.64)	1.539^{c}	(11.3)	1.432^{c}	(8.87)
Bullishness B^{*}_{it}	1.506^{a}	(2.54)	2.149^{a}	(2.53)	1.937^{c}	(4.33)			3.950^{c}	(4.67)
Agreement A_{it}			0.485	(1.38)			1.635	(1.92)	0.226^{b}	(2.80)
Log likelihood	−3691		−3690		−3712		−3720		−3708	
Observations	9828		9828		9828		9828		9828	
Companies	39		39		39		39		39	

	Does Message Posting Today Predict Articles in the WSJ Tomorrow?									
	(1)		(2)		(3)		(4)		(5)	
Messages $\ln(1+M_{it})$	1.635^{c}	(10.5)	1.571^{c}	(9.23)						
Words $\ln(1+W_{it})$					1.303^{c}	(7.60)	1.384^{c}	(9.55)	1.262^{c}	(6.60)
Bullishness B^{*}_{it}	2.051^{c}	(4.63)	4.065^{c}	(4.78)	2.628^{c}	(6.53)			7.291^{c}	(6.95)
Agreement A_{it}			0.264^{b}	(2.72)			1.829^{b}	(2.62)	0.131^{c}	(4.11)
Log likelihood	−3696		−3692		−3722		−3740		−3712	
Observations	9828		9828		9828		9828		9828	
Companies	39		39		39		39		39	

	Does Message Posting Today Predict Articles in the WSJ the Day after Tomorrow?									
	(1)		(2)		(3)		(4)		(5)	
Messages $\ln(1+M_{it})$	1.352^{c}	(6.94)	1.321^{c}	(6.16)						
Words $\ln(1+W_{it})$					1.148^{c}	(4.77)	1.170^{c}	(5.56)	1.132^{c}	(4.25)
Bullishness B^{*}_{it}	1.160	(.974)	1.748^{a}	(2.11)	1.362^{a}	(2.10)			2.524^{c}	(3.62)
Agreement A_{it}			0.458	(1.90)			0.932	(.311)	0.300^{b}	(2.92)
Log likelihood	−3770		−3768		−3782		−3785		−3778	
Observations	9828		9828		9828		9828		9828	
Companies	39		39		39		39		39	

VII. Conclusions

This paper studies how Internet stock message boards are related to stock markets. We find that there is useful information present on the stock message boards. The magnitudes of some of the observed effects are quite large relative to other features of stock markets that have attracted attention. Our main specific questions are as follows. Do the messages successfully predict subsequent stock returns? Is disagreement among the messages associated with increased trading volume? Do the messages help to predict subsequent volatility?

First, we find that a positive shock to message board posting predicts negative returns on the next day. This effect is statistically significant but economically quite small in comparison to plausible transactions costs. This negative return on the next day has not been previously reported in the literature.

Second, a traditional hypothesis is that disagreement induces trading. We find significant evidence supporting this claim in contemporaneous regressions. Both the level of message posting and disagreement among the messages helps to predict subsequent trading volume – particularly for smaller-sized trades. However, greater disagreement on one day predicts fewer trades on the next day, not more trades.

Third, message posting does help to predict volatility both at daily frequencies and also within the trading day. This effect is found within a range of methods of modeling market volatility, including realized volatility methods and GARCH class methods. Message posting activity is a readily observed correlate that helps to account for volatility.

The stock messages reflect public information extremely rapidly. Accordingly, Internet stock messages may be helpful in studies of insider trading and event studies. Since the messages are time-stamped to the minute, they may also prove quite helpful in market microstructure studies.

The evidence clearly rejects the hypothesis that all that talk is just noise. There is financially relevant information present. In some respects the information even goes beyond what can be found in the WSJ. The talk has only a very minor effect on high-frequency stock returns—an effect that seems to be swamped by plausible transactions costs. The talk appears to be very pertinent for predicting trading volume and volatility.

REFERENCES

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001, The distribution of stock return volatility, *Journal of Financial Economics* 61, 43–76.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2001, The distribution of realized exchange rate volatility, *Journal of the American Statistical Association* 96, 42–55.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2002, Modeling and forecasting realized volatility, Working paper, University of Pennsylvania.

Bacidore, Jeffrey Michael, and Marc L. Lipson, 2001, The effects of opening and closing procedures on the NYSE and Nasdaq, Working paper, NYSE and University of Georgia.

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

Bessembinder, Hendrik, Kalok Chan, and Paul J. Seguin, 1996, An empirical examination of information, differences of opinion, and trading activity, *Journal of Financial Economics* 40, 105–134.

Cao, Henry H., Joshua D. Coval, and David Hirshleifer, 2002, Sidelined investors, trade-generated news, and security returns, *Review of Financial Studies* 15, 615–648.

Chan, Kalok, and Wai-Ming Fong, 2000, Trade size, order imbalance, and the volatility-volume relation, *Journal of Financial Economics* 57, 247–273.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2001, Market liquidity and trading activity, *Journal of Finance* 56, 501–530.

Cover, Thomas M., and Joy A. Thomas, 1991, *Elements of Information Theory* (Wiley, New York).

Das, Sanjiv R., and Mike Chen, 2001, Yahoo! for Amazon: Sentiment parsing from small talk on the web, Working paper, Santa Clara University.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel, 2001, Persuasion bias and social influence, Working paper, Stanford University.

Dewally, Michael, 2000, Internet investment advice: Investing with a rock of salt, Working paper, University of Oklahoma.

Diamond, Douglas W., and Robert E. Verrecchia, 1981, Information aggregation in a noisy rational expectations economy, *Journal of Financial Economics* 9, 221–235.

Engle, Robert F., and Andrew J. Patton, 2001, What good is a volatility model? *Quantitative Finance* 1, 237–245.

French, Kenneth R., G. Willaim Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Geweke, John, and Susan Porter-Hudak, 1983, The estimation and application of long-memory time-series models, *Journal of Time Series Analysis*, 4, 221–238.

Glosten, Lawrence, Ravi Jagannathan, and David Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

Hamilton, James D., 1994, *Time Series Analysis* (Princeton University Press, Princeton, NJ).

Hansen, Peter Reinhard, and Asger Lunde, 2001, A forecast comparison of volatility models: Does anything beat a GARCH (1,1)? Working paper, Brown University.

Harris, Milton, and Artur Raviv, 1993, Differences of opinion make a horse race, *Review of Financial Studies* 6, 473–506.

Hirshleifer, Jack, 1977, The theory of speculation under alternative regimes of markets, *Journal of Finance* 32, 975–999.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2002, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, Working paper, Stanford University.

Jones, Charles M., Gautam Kaul, and Marc L. Lipson, 1994, Transactions, volume, and volatility, *Review of Financial Studies*, 7, 631–651.

Kandel, Eugene, and Neil D. Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831–872.

Karpoff, Jonathan M., 1986, A theory of trading volume, *Journal of Finance* 41, 1069–1088.

Lakonishok, Josef, and Edwin Maberly, 1990, The weekend effect: Trading patterns of individual and institutional investors, *Journal of Finance* 45, 231–243.

Lewis, David D., 1998, Naive (Bayes) at forty: The independence assumption in information retrieval, in *Proceedings of the 10th European Conference on Machine Learning (ECML)* (Springer-Verlag, New York).

Manning, Christopher D., and Hinrich Schütze, 1999, *Foundations of Statistical Natural Language Processing* (MIT Press, Cambridge, MA).

Marcinkowski, Victoria, 2001, Skechers USA up 9%; buyers step in after FMR gets out, *Dow Jones News Wires*, August 14.

McCallum, Andrew, 1996, Bow: A toolkit for statistical language modeling, text retrieval, classification and clustering, Working paper, School of Computer Science, Carnegie-Mellon University.

Miles, Stephanie, 2001, Flooz.com says it is seeking merger, but service remain offline for now, *Wall Street Journal, Online Edition*, August 10.

Milgrom, Paul, and Nancy L. Stokey, 1982, Information, trade and common knowledge, *Journal of Economic Theory* 26, 17–27.

Schwert, G. William, 1990, Stock volatility and the crash of '87, *Review of Financial Studies* 3, 77–102.

Shiller, Robert J., 2000, Conversation, Information, and Herd Behavior, *American Economic Review* 85, 181–185.

Tumarkin, Robert, and Robert F. Whitelaw, 2001, News or noise? Internet message board activity and stock prices, *Financial Analysts Journal* 57, 41–51.

Weiss, Garry, 2000, Where everyone is an expert, *BusinessWeek* May 22, 173–176.

Wysocki, Peter D., 1999, Cheap talk on the web: The determinants of postings on stock message boards, Working paper, University of Michigan.