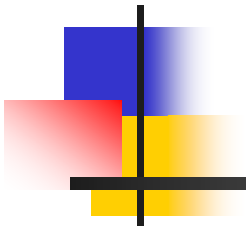
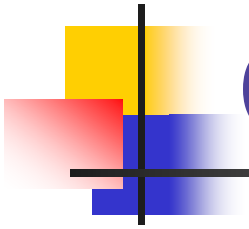


Genetic Programming and the Predictive Power of Internet Message Traffic

James D Thomas

Katia Sycara

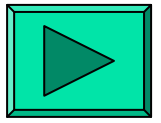


Outline

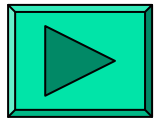
- Introduction
- Data
- Trading Rules Framework
- Measures of Success
- A GP Learner
- Empirical Results
- Summary

Introduction

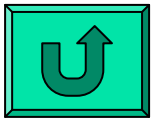
- Uses genetic algorithms to examine the relevance of one new source of information -- the volume of message board postings on stock specific message boards on the financial discussion areas of yahoo.com and ragingbull.com.

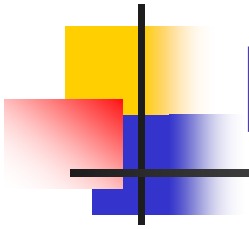


- The key question is if the measures of message volume can be used as an effective predictor of stock movements.
- They build a specialized GP learner that builds trading rules based on this message volume data.



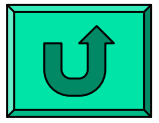
- They have performed preliminary explorations on smaller versions of this data set. (Thomas and Sycara, 2000).
- This paper extends those techniques to a larger datasets, generating more robust conclusions.





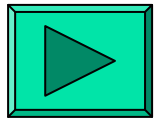
Data

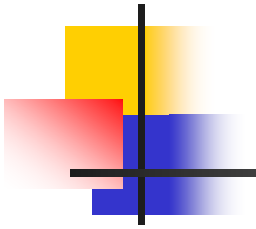
- Select Stocks
- Time Universe
- Split the Set of Stocks in Half
- Market Data
- Message Traffic Data



Select Stocks

- They limited the universe of stocks were those that appeared on the **Russell 1000** (a list of the 1000 largest US equities by market capitalization, updated yearly) index for both 1999 and 2000, and who had price data dating back to Jan 1, 1998, on the yahoo.com quote server. This left us with 688 stocks.

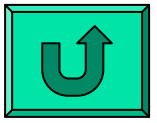




- we limited ourselves to the top 10% by *message traffic volume*, leaving us with 68 stocks.

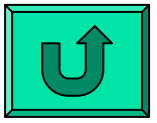
Time Universe

- January 1, 1998 to December 31, 2001.



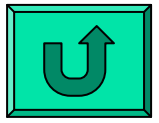
Split the Set of Stocks in Half

- Randomly split this set of stocks in half
  - One half is used as a design set to build the algorithm.
  - The other half is used as a holdout test set to verify the results.



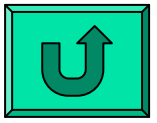
Market Data

- Downloaded split adjusted prices and trading volume off of the yahoo.com quote server for each stock.
- Use those price figures to compute excess returns.
- We realize that this ignores dividends and renders the excess return figures inexact; however, since most of the bulletin board with high discussion are technology companies who pay no dividends, we feel that this is an acceptable compromise.



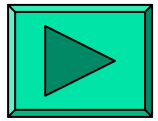
Message Traffic Data

- For the message traffic data itself, we collected posts off of both the yahoo.com and ragingbull.com bulletin boards for every stock in the stock universe.
- Handle these counts of message board volume

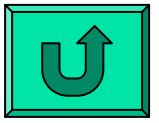


Handle These Counts of Message Board Volume

- Only posts made while markets were closed were counted. (Information contained in posts made during market open should be factored quickly into the prices.)
- The daily count of messages was normalized by a factor determined by the day of the week, so that the expected number of posts on each day of the week was the same.

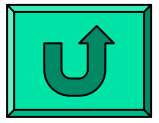


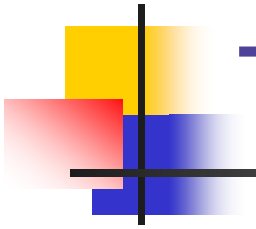
- For multi-day periods when the markets were closed (weekends or holidays), message counts for the appropriate non-market days were averaged.
- We added the message traffic volume from ragingbull.com and yahoo.com together to get a single message count.



Trading Rules Framework

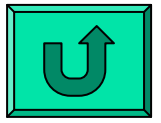
- Task
- Make a Decision
- Definitions
- The Formula for Daily log Returns
- Fitness measure : returns
  - Maximize the total returns
  - Not Maximize prediction accuracy





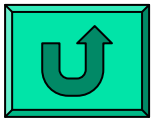
Task

- To learn trading rules over a universe of stocks that perform better than merely buying and holding the universe of stocks.



Make a Decision

- For each stock, we make a basic decision: long, or short.
- If we decide to short a stock, we take a corresponding long position in the broader market (proxied by the Russell 1000 index).



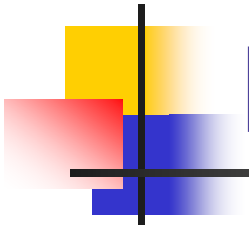
Definitions

- Let $r_{Strategy}$ be daily log return our strategy produces
- Let x(t) be our trading signal: 1 for 'long', 0 for 'short'.
- Let $r_{stock}(t)$ be the daily log return on the stock at time t
- Let $r_{Russell1000}$ (t) be the daily log return on the Russell 1000 at time t
- Let tcost be the one-way log transaction cost.
- Let $r_{shortrate}$ be the rate we pay

?

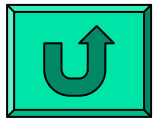
The Formula for Daily log Returns

- if $x(t) = 1, r_{Strategy}t = r_{stock}(t) - \left|x(t) - x(t-1)\right| \cdot tcost$
- if $x(t) = 0, r_{Strategy}(t) = r_{Russell1000} - r_{stock}(t) - \left|x(t) - x(t-1)\right| \cdot tcost - r_{shortrate}$



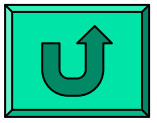
Measures of Success

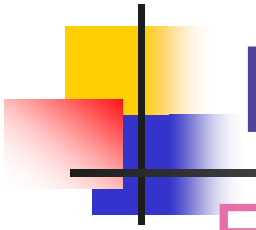
- Benchmark
- Performance
- Significance
- Avoid Overfitting



Benchmark

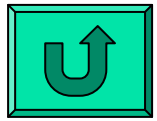
- Buy and hold strategy over the appropriate stocks
- If our trading strategy can produce risk adjusted excess returns while accounting for reasonable transaction costs, then this is a strong argument that the algorithm is picking up a meaningful pattern in the data.

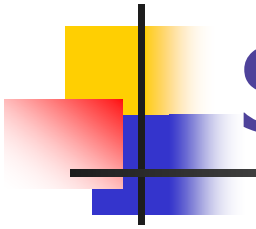




Performance

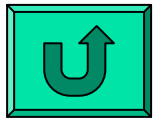
- Excess Returns
- Excess Sharpe Ratio
  - The Sharpe ratio of the trading strategy minus the Sharpe ratio of the buy and hold strategy, where both Sharpe ratios are computed against the an assumed risk free rate of 5%.
- Sharpe Ratio
  - The Sharpe ratio of the trading strategy against a benchmark of the buy-and-hold strategy.





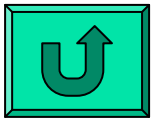
Significance

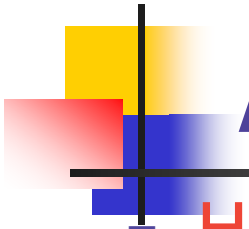
- Bootstrap hypothesis testing
  - Define the null hypothesis.
  - Generate a number of datasets by the null hypothesis.
  - Run the algorithm on these bootstrap datasets.
  - Compare what proportion of the bootstrap datasets produce results exceeding that of the real dataset; this is the appropriate p-value.



Null Hypothesis

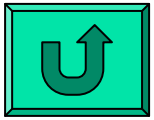
- The message volume statistics associated with a trading day has no predictive power.

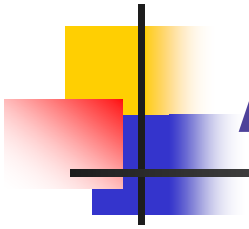




Avoid Overfitting

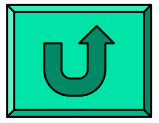
- Hold out a final testing set of data. This data will not be touched until the algorithm design process is complete.
- Split the remaining data into training and testing sets.
- Perform algorithm design on only this data -- develop the algorithm by examining performance on the test set.
- Then, only when the algorithm has been settled, verify the conclusions based on the "holdout" set.





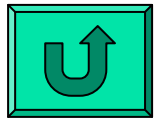
A GP Learner

- GP
  - Basic Algorithm
  - Parameters
- Relearn Periodically
- Representation

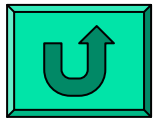


Basic Algorithm (no crossover)

- Split data into training, validation, and testing set.
- Generate a random population of trading rules.
- Run the following algorithm for n generations.
  - Evaluate the fitness of the entire population.
  - Perform selection and create a new population.
  - Mutate the surviving population.
- After this training phase is over, take the final population, and select the trading rule with the highest fitness on the validation set.
- Evaluate this individual's fitness on the testing set.

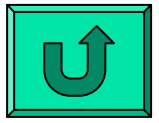


- The training and validation sets are always a 50/50 split of the available training data.



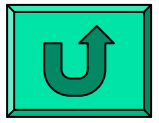
Parameters

- Population size : 20
- Generations : 10
- Selection :
  - Binary deterministic tournament : Two distinct individuals selected randomly with uniform probability compete at each tournament.
- Fitness : Returns
- Maximum number of nodes : 10



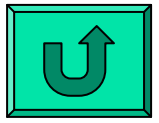
Relearn Periodically

- To avoid applying trading rules to a data in test set temporally distant from the training set.
- Start :
  - Training/validation set (split 50/50) : **1998.1—1998.6**
  - Test set : **1998.7—1998.9**
- Then :
  - Training/validation set (split 50/50) : **1998.1—1998.9**
  - Test set : **1998.10—1998.12**



Representation

- Past work :
  - "in" or "out" of the asset with roughly equal probability.
  - Implicit Assumption : every day is equally easy for the learner to predict.
- If the current message traffic volume is greater than a threshold, we get out of the stock, and stay out for a period of time.
  - We do not always want to make a prediction.
  - We only care about spikes in message volume traffic.
- Format



Format

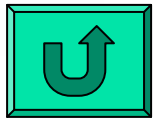
- if $v_t > ma_{t,n} + k\sigma_{t,n}$, sell the stock short and go long in the market.

Where:

- v_t is the message volume from market close of day $t-1$ until market open on day t.
- $ma_{t,n}$ is the n day moving average over message volume up to day t
- $\sigma_{t,n}$ is the n day standard deviation of the message volume up to day t
- m is the length of time we stay out of the stock once we've detected an event.

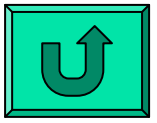
?

- The ranges of the parameters



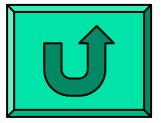
The Ranges of the Parameters

- n, mean and standard deviation window size: 10 to 100 trading days, in increments of 10 trading days
- k, event threshold, measured in standard deviations: 3 to 6 standard deviations above the mean, in increments of .5.
- m, stay out period: One to ten trading days.



Empirical Results

- The Standard Approach
- Other Possible Predictive Variables
- Changing the Nature of the Trading Rules
- Test on Holdout Data
- Regime Changes

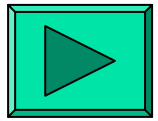


The Standard Approach

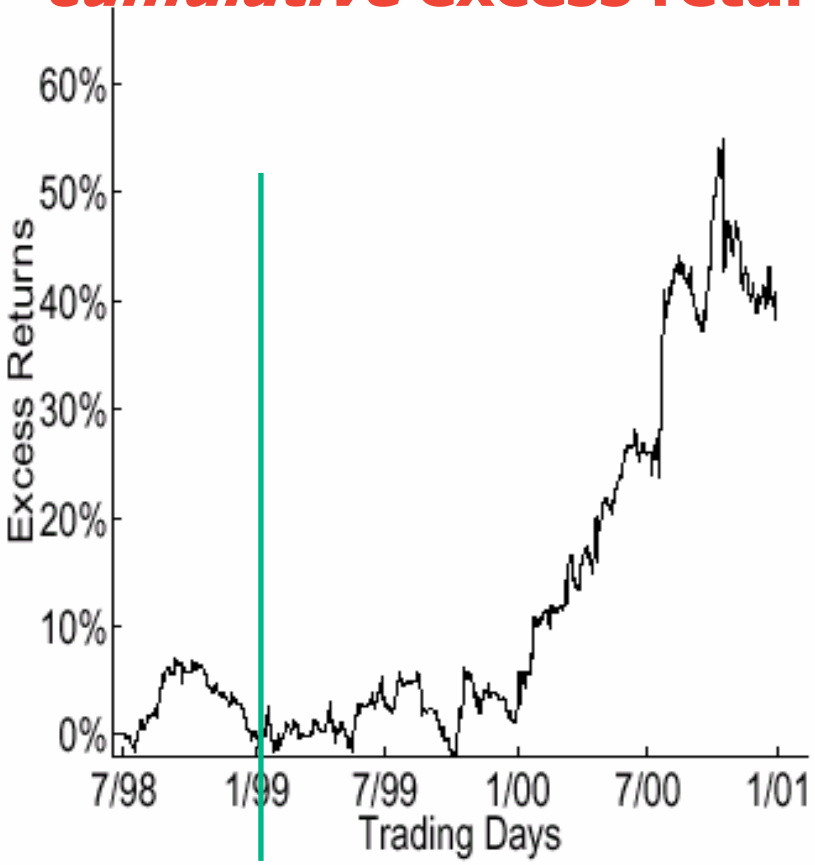
- 200 bootstrap datasets
- 30 trials

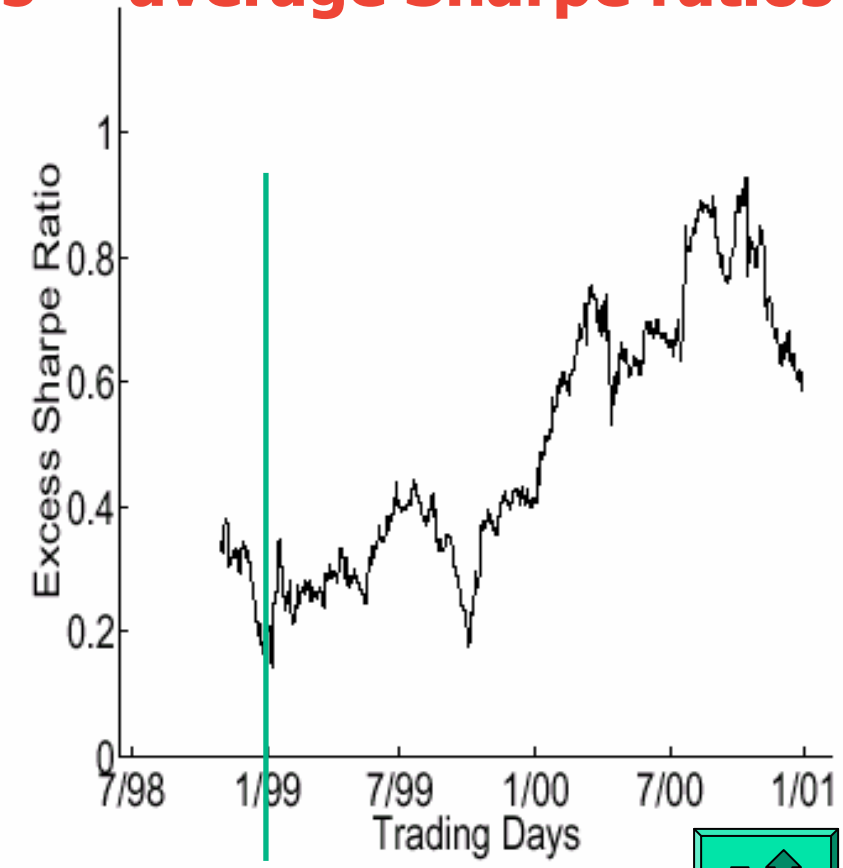
Approach:	GP Learner	Buy and Hold	Difference	Bootstrap p
Returns	164.56%	126.21%	38.34%	.015
Daily Std Dev	2.18%	2.45%	.27%	.010
Sharpe Ratio	1.7455	1.1598	.5857	.005
Differential Sharpe	N/A	N/A	1.6076	.030

??



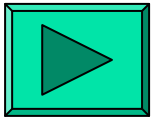
"*cumulative* excess returns" "average Sharpe ratios"





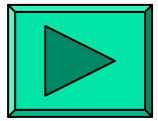
Other Possible Predictive Variables

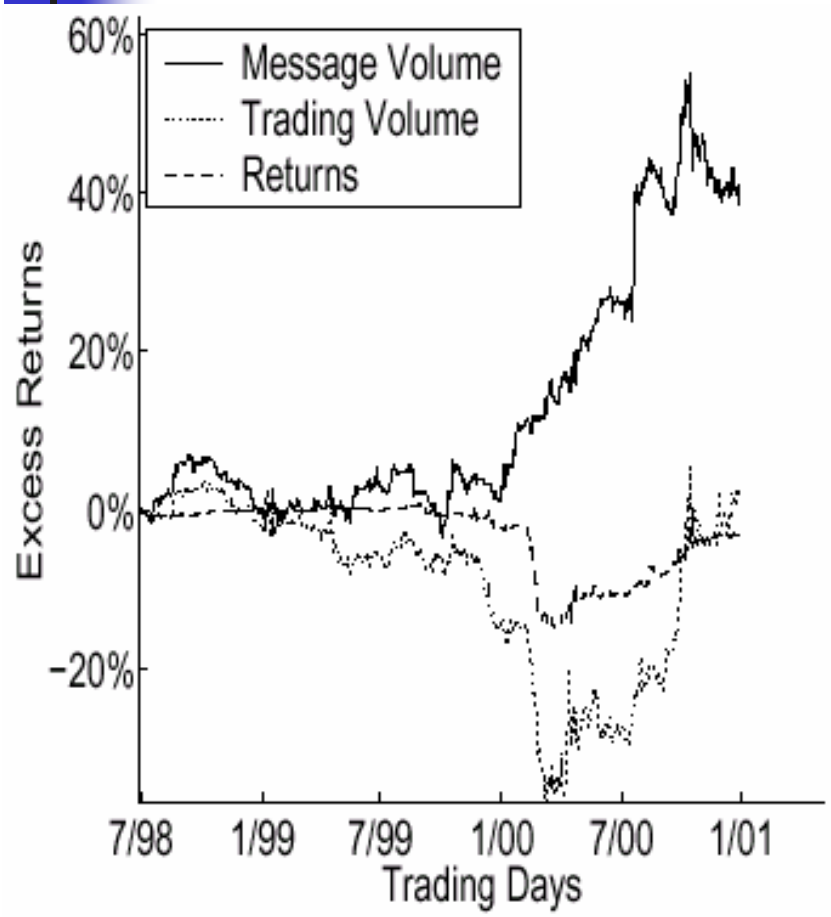
- There is some correlation between message traffic volume and other variables
  - r(lagged trading volume, message traffic)= .5194
    - The high correlation between message volume and trading volume suggests the possibility that message volume is simply echoing trading volume.
  - r(lagged returns, message traffic)= -.1017.
    - Lagged returns are unlikely to contain the same information as the message volume.

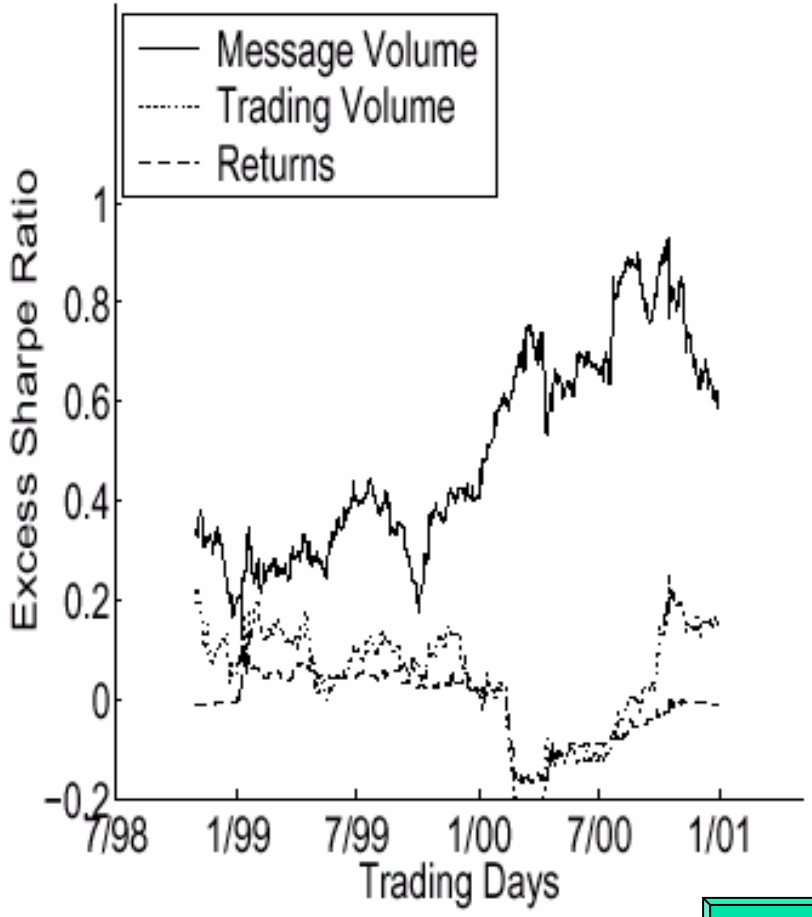


Predictor Variable:	Excess Returns	Excess Sharpe	Differential Sharpe
Message Volume	38_34%	_5857	1_6076
Trading Volume	1_27%	_1519	_0438
Returns	3_13%	__0088	__3569

- Using a 2-tailed T test we found that the differences between the message volume results and the lagged trading volume and lagged returns results were all statistically significant, with p-values less than .001 in all cases.




Message Volume
Trading Volume
Returns
60%
40%
20%
0%
−20%
Excess Returns
7/98
1/99
7/99
1/00
7/00
1/01
Trading Days


Message Volume
Trading Volume
Returns
1
0.8
0.6
0.4
0.2
0
−0.2
Excess Sharpe Ratio
7/98
1/99
7/99
1/00
7/00
1/01
Trading Days

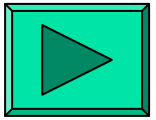
Changing the Nature of the Trading Rules

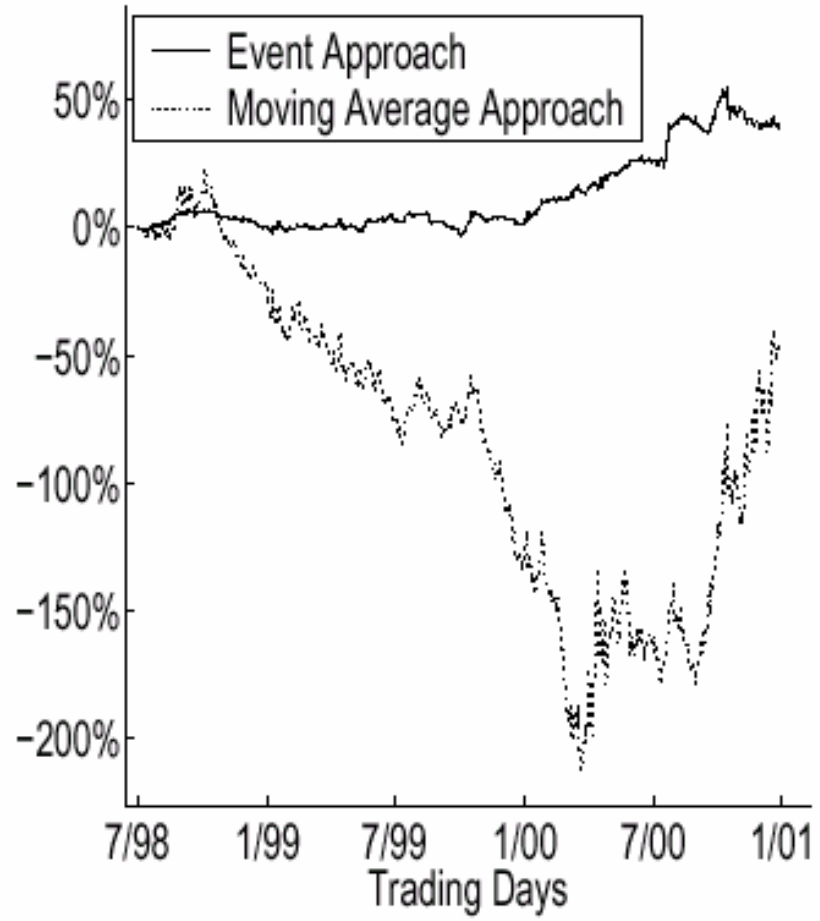
- Key difference: instead of looking for a rare event and pulling out of a stock, this kind of trading rule is neutral with regards to being in or out of a stock.

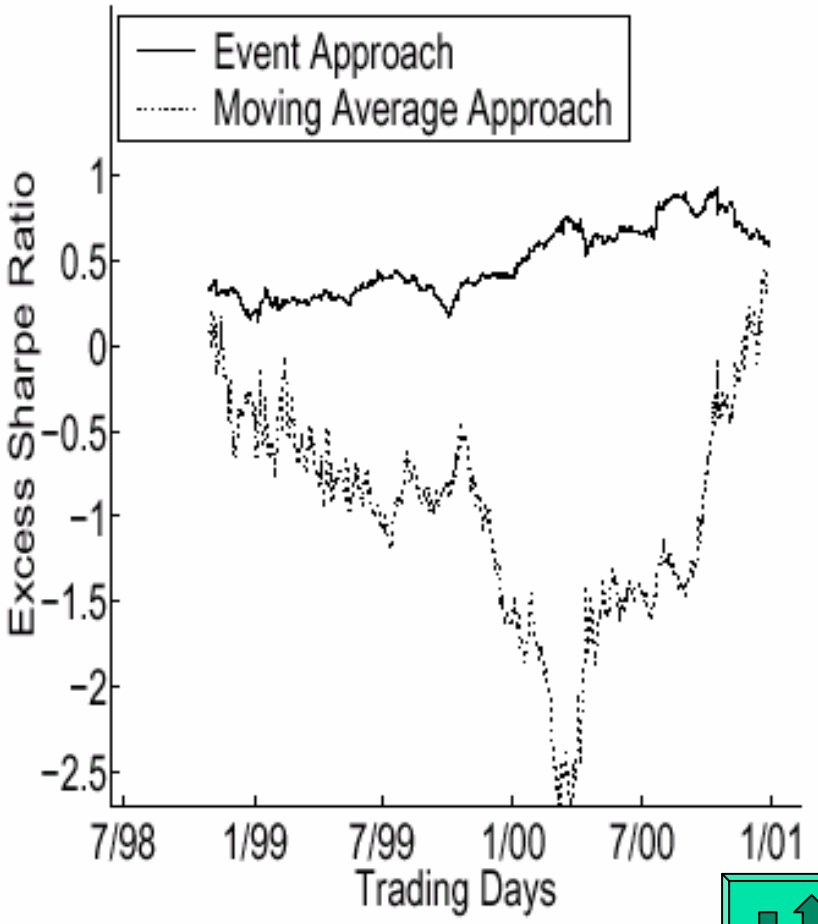
- if $ma_{t,n} > ma_{t,m}$, sell the stock short and go long in the market.

Trading Rule Scheme:	Excess Returns	Excess Sharpe	Differential Sharpe
Event approach	38.34%	.5857	1.6076
Moving average approach	-46.39%	-.4177	-.4751

The volatility of the moving average approach is very low.



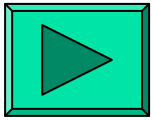

Event Approach
Moving Average Approach
50%
0%
−50%
−100%
−150%
−200%
7/98
1/99
7/99
1/00
7/00
1/01
Trading Days

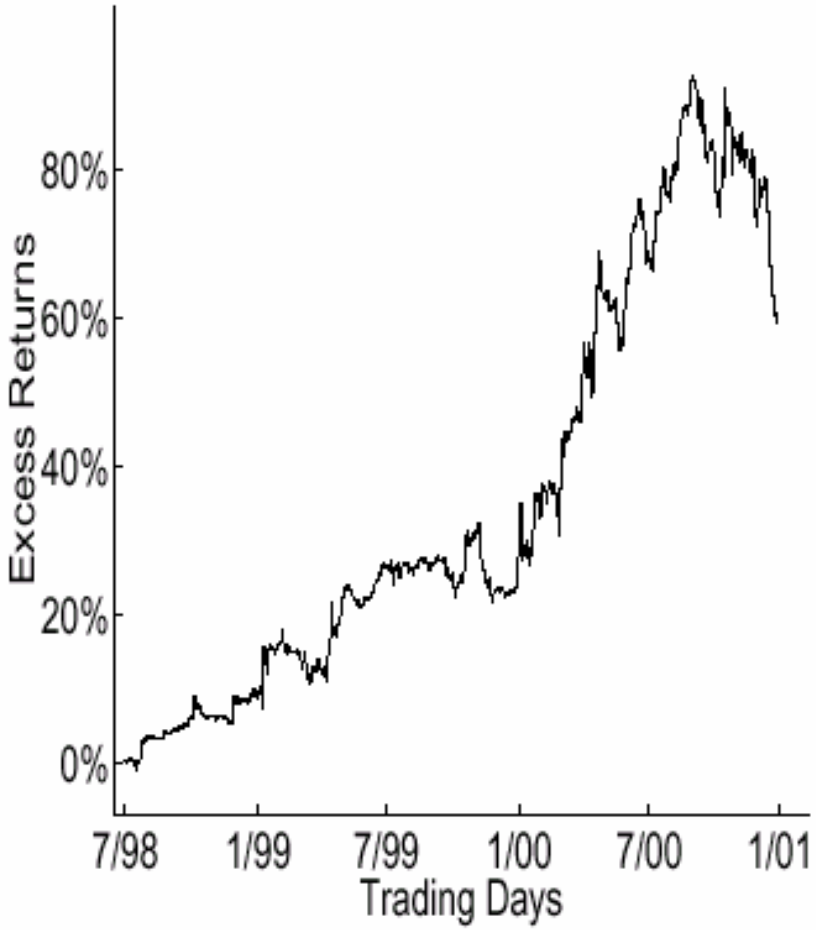

Event Approach
Moving Average Approach
Excess Sharpe Ratio
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
7/98
1/99
7/99
1/00
7/00
1/01
Trading Days

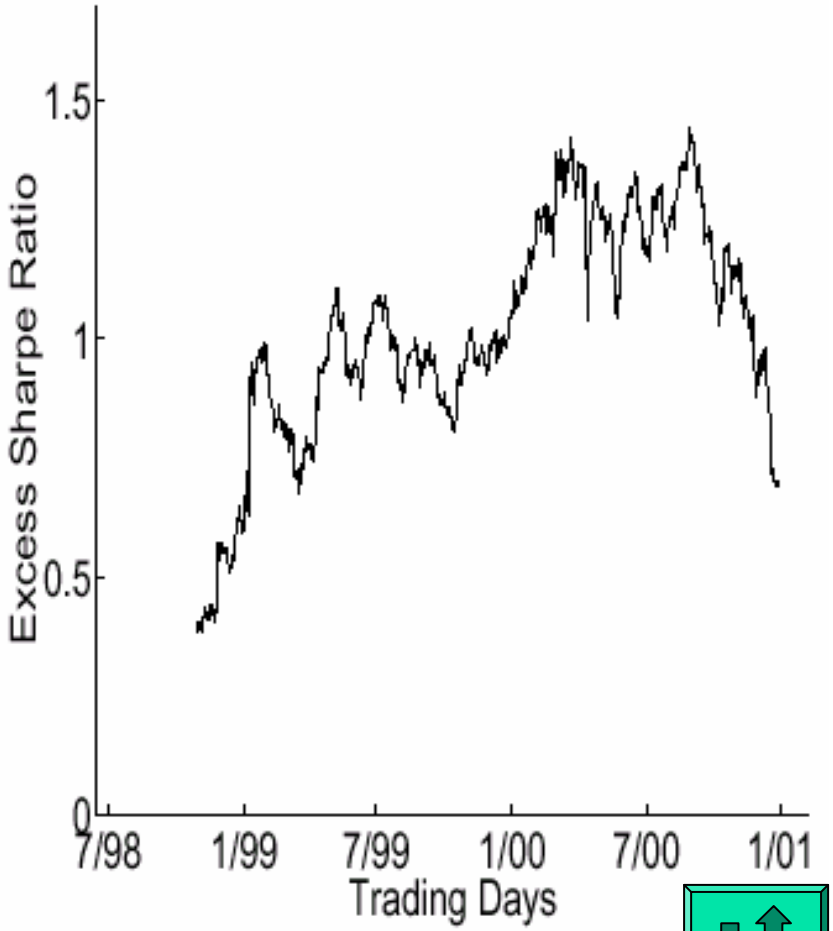
Test on Holdout Data

Approach:	GP Learner	Buy and Hold	Difference	Bootstrap p
Returns	149.01%	89.60%	59.41%	.045
Daily Std Dev	2.59%	3.03%	-.44%	.005
Sharpe Ratio	1.3235	.6353	.6882	.030
Differential Sharpe	N/A	N/A	1.7548	.110

- The p-values are higher than in the test set.
- The excess returns and excess Sharpe ratio are still statistically significant by the bootstrap hypothesis testing.

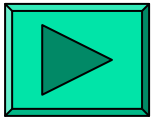



Excess Returns
80%
60%
40%
20%
0%
7/98
1/99
7/99
1/00
7/00
1/01
Trading Days


Excess Sharpe Ratio
1.5
1
0.5
0
7/98
1/99
7/99
1/00
7/00
1/01
Trading Days

Regime Changes

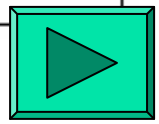
- Excess returns decline on both the test set and the holdout data set from October of 2000 to the end of the time period.
- Will it continue?
- Instead of looking for spikes in message volume, we look for **slumps** in message volume.

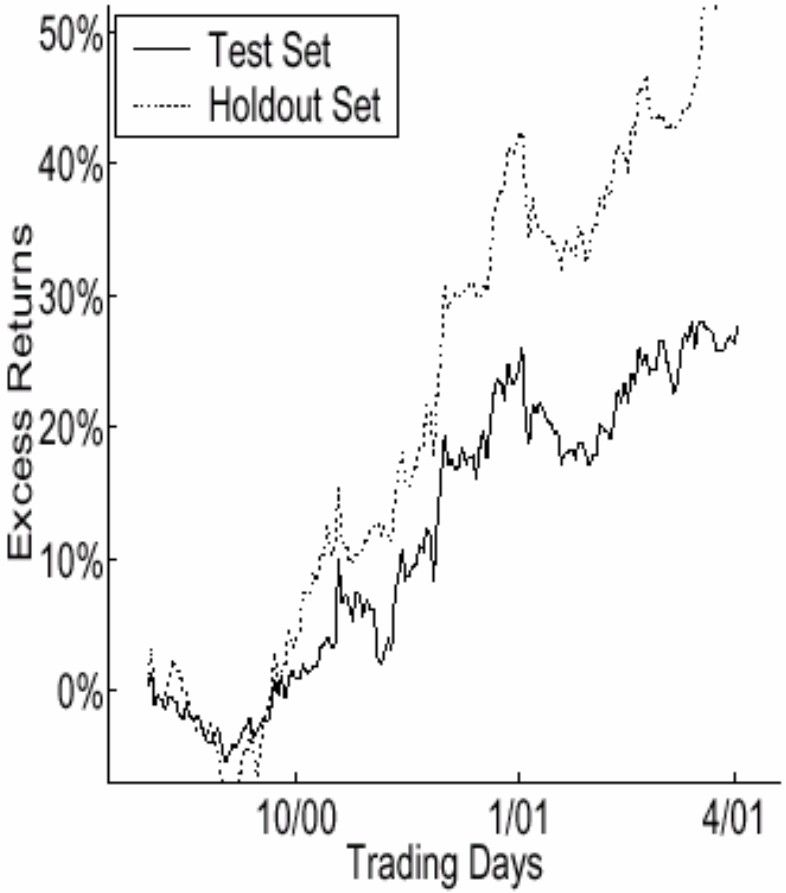


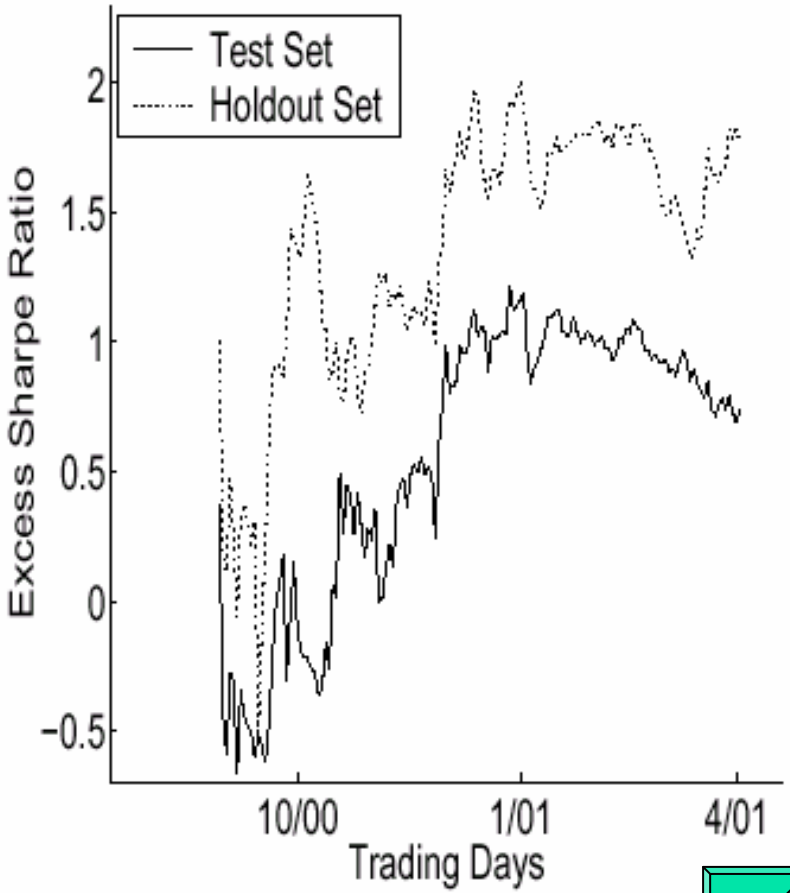
- if $v_t < ma_{t,n} + k\sigma_{t,n}$, sell the stock short and go long in the market.

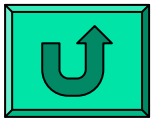
- change the range of minimum event thresholds from 3 to 6, to -1.5 to -3, and search in increments of .25. (The distribution of message volume traffic is skewed.)

Approach:	GP Learner	Buy and Hold	Difference	Bootstrap p
Performance on Test Set				
Returns	-1.03%	-28.59%	27.56%	.020
Daily Std Dev	1.64%	3.07%	-1.43%	.005
Sharpe Ratio	-.2390	-.9817	.7427	.020
Differential Sharpe	N/A	N/A	1.2640	.235
Performance on Holdout Set				
Returns	21.15%	-35.94%	57.09%	.005
Daily Std Dev	2.10%	3.88%	-1.78%	.005
Sharpe Ratio	.8151	-.9549	1.7701	.005
Differential Sharpe	N/A	N/A	2.1855	.025




Test Set
Holdout Set
50%
40%
30%
20%
10%
0%
Excess Returns
10/00
1/01
4/01
Trading Days


Test Set
Holdout Set
2
1.5
1
0.5
0
−0.5
Excess Sharpe Ratio
10/00
1/01
4/01
Trading Days

Summary

- The message board volume data has predictive power.
- The message board volume data contributes information that other traditional numerical data (price, volume, etc) are not.