
SHARE PRICE TREND BASED ON ANALYSIS OF INVESTOR'S ONLINE ACTIVITY¹

PhD Univ. Lecturer Mihail Dumitru SACALĂ
Academy of Economic Studies, Bucharest

Abstract

The desire of modeling the influence coming from behavioral economics, representing mostly qualitative variables, requires new approaches to quantify this influence. In this paper we present an analytical method for determining the direction of the share future return using an aggregate indicator calculated based exclusively on the trend of registered variables from investor's online activity. This online activity refers to the number and content of messages posted by investors on shares analysis dedicated web blogs.

Key words: asset pricing, behavioral finance, qualitative variables, investors.

Often the future evolution of prices and trading volume for a share are highlighted through the online activity of investors. This activity can be captured in one of two situations:

1. **Analysis of listed companies** search through search engines or social networking sites (Google, Twitter, Facebook, etc.). Thus, in 2010, Bollen et al., determined a forecasting model for U.S. capital market indexes based on twitter users activity.

Tobias Preis (November 15, 2010) analyzed the searches on Google for S & P 500 component companies along with their price and volume. He failed to determine the weekly price fluctuations but found a strong correlation between the number of Internet searches (Google Trends) and the trading volume for the company's shares.

An academic group access in Google databases is required in order to predict shares price trend due to investor perceptions. Stock market crash can be also prior determined when we have information on the number of searches in each hour for each stock.

2. **The modifications of the comments made by investors** upon capital market news and papers. The determined indicators examine the

¹ This paper was co-financed by European Social Fund, the Human Resources Development Operational Program 2007-2013, the number POSDRU/89/1.5/S/59184 "Performance and excellence in postdoctoral research in economic sciences in Romania."

frequency of these messages and votes obtained by each of them from other participants in online discussion.

This paper studies the second situation. The analyzed messages are representative samples for the opinions of investors for the specific company. Active people on blog and other electronic media are those that invest for short and medium term, have a higher average number of transactions than other investors and thus they affect trading volume and price.

Investors interested in a particular stock, wanting to obtain as much information about that company, read articles on the Internet using both stock exchanges exclusive blogs and electronic publications.

Many of these Web sites allow users to post messages and comments for articles that refer to a certain company.

These messages fall into two categories: Comments related to an article; Responses to another post.

This paper presents a methodology for calculating the index and its interpretation for each of the two situations described above.

Analyzing the variables identified by studying these kinds of messages we can determine an indicator that will give us signals about how the price and the trading volume for a specific company's shares will be modified in the immediately following period.

The intensity of online activity is closely linked to investor's interest for that instrument. A lot of decisions to buy or sell shares are taken from the interpretation of these messages. This causes both a change in the volume traded and the trading price.

After their contents, these messages fall into two categories:

- **Messages which show an expected** effect of short-term yields (these messages come from one of the following situations: 1. the investor present his view on the share's trend, 2. hopes that this movement will be achieved, 3. thinks that this message can influence the market in its favor). These messages are posted by people who want to trade this action in the near future.

- **Neutral messages** (these messages are not related to that company or they do not present a point of view for the discussed share price movement).

The variables we may identify vary depending on how the blog, web page etc is built. Theoretical, the variables which may be analyzed are:

- **The message date** - is the date when the message was posted as a comment to an article or on a dedicated page. This variable shows the distance in days between the time when the article was posted and the time when this message was posted. The greater the distance is the greater impact of the article will be and therefore the indicator value for that article online activity will be higher.

- **The number of responses** received from investors by every posted message. If the number of responses is high, the message has an increased impact on the perceptions of investors and therefore it will lead to an increase of the online activity indicator for the studied paper.

- **The number of votes** received by each users posted message. A large number of votes involve a huge interest in the message.

- **The grade for each message** (the difference between the number of affirmative votes and the number of negative votes). A low grade in terms of a large number of votes indicates an increased activity.

- **The sign of planned** returns for specified share, subtracted from the message.

Neighborhoods of extreme points for a stock price

The papers which discuss a particular stock appear in various electronic publications without having continuity. In this particular situation we can not build an indicator to characterize continuously the online activity for this company. The only method of analysis is to build an indicator, the values of which will classify the papers and therefore to identify changes that will occur in the price and trading volume for shares of the company.

For example we used www.hotnews.ro news portal. Dates of occurrence of the analyzed four items are: December 19, 2007; July 10, 2008; April 27, 2009; March 2, 2011

Appendix 1 provides an example of message analysis. The message appeared as a comment on one of the analyzed four items.

“Records FP (Wednesday, March 2, 2011, 6:28 p.m.)

Nae [anonymous]

A record for FPR volume, but the price is still low due to bad company contract made by ... “non-owners, obviously. “

I_{articol} is the indicator which characterizes the activity of investors in a paper. Based on the values of this indicator we may determine the following changes for the trading volume and the share price.

Proposed formula for this indicator is:

$$I_{\text{articol}} = \frac{\sigma_{\text{sens}}^2 * \sum_{i=1}^n [d_{ci} - (d_a - 1)](n_{\text{raspi}} + 1) * k_i * \left(\frac{nvot_i + 1}{nota_i + 1}\right)}{\sum_{i=1}^n (nvot_i + 1)};$$

Where:

n - the total number of messages recorded for the current paper;
 σ_{sens}^2 dispersion of the binary variable: “forecasted sign of the future return”.

$$\sigma_{\text{sens}}^2 = \frac{n_+ * n_-}{n_{\text{sens}}^2};$$

Where: n_+ - the number of messages that predict an increase of the price of the stock;

n_- - the number of messages that predict a decrease of the price of the stock;

n_{sens} - the number of messages that predict the sign for future returns of the specified stock.

d_{ci} - the date when the “i” message was posted;

d_a - the date when the discussed paper appeared;

n_{raspi} - the number of responses for the “i” message;

k - the factor that highlights the importance of forecasting the direction of the future return:

$k= 2$ if the message predicts the sign for the future performance;

$k= 1$ otherwise;

$nvot_i$ - the number of votes obtained by the “i” message;

$nota_i$ - the grade obtained by the “i” message.

Given the methodology of defining the indicator $I_{\text{articol}} \in [0; +\infty)$, and its values may fall an article in the following categories:

$I_{\text{articol}} \in [0; 0,3]$, the online activity is low, the trading volume will be low, while the sign of return can not be established due to high volatility.

$I_{\text{articol}} \in (0,3;1]$, the online activity is moderate, the trading volume will fluctuate around the average and the price will follow the short-term trend.

$I_{\text{articol}} \in (1; +\infty)$, the online activity is high, we are in the neighborhood of an extreme point for the price trend and this indicates that the sign for the return on short term will be opposite to the present average return. We expect a high trading volume.

Observed values for the variables identified in an article

Comment Date	Number of answers	Number of votes	Vote result	Foresight	dc-da+1	Number of answers +1	K
27	4	9	7	s	1	5	2
27	0	3	1	s	1	1	2
27	2	1	1	c	1	3	2
27	0	5	-1	s	1	1	2
28	0	0	0	c	2	1	2
27	2	4	2	c	1	3	2
27	0	0	0	n	1	1	1
27	0	3	3	s	1	1	2
28	0	0	0	s	2	1	2
27	0	4	2	s	1	1	2
27	0	7	5	s	1	1	2
27	0	5	3	s	1	1	2
27	1	1	1	c	1	2	2
27	0	1	-1	s	1	1	2
27	3	3	-1	n	1	4	1
27	1	0	0	n	1	2	1
27	0	1	1	n	1	1	1
27	0	2	0	s	1	1	2
27	2	5	-1	s	1	3	2
27	1	0	0	n	1	2	1
27	0	1	1	n	1	1	1
27	1	6	0	c	1	2	2
27	0	2	2	s	1	1	2
27	4	0	0	s	1	5	2
27	3	1	1	c	1	4	2
27	2	0	0	c	1	3	2
27	0	0	0	s	1	1	2
28	0	0	0	s	2	1	2
29	0	0	0	n	3	1	1

Appendix 2 shows the recorded values for the identified variables for each of the four articles. The values of indicators in the four cases are: 1. 0,99; 2. 1,33; 3. 0,79; 4. 1,6

In online, dedicated to stock exchanges, forums (eg www.agf.ro Blogs General Assembly) messages related to certain companies appear every day. In this situation it is necessary to build an indicator that can be calculated daily and therefore online activity can be characterized at all times.

Example of message analysis, recorded on www.agf.ro portal

24-May-2011 4:23 p.m. - stef_nic: if the stock market next year remains at this size 0,2 RON is likely for fp

The message appeared as a comment on the forum page dedicated to FP stock.

Proposed indicator has the following formula:

$$I_t = \frac{n_t}{\frac{\sum_{i=t-19}^t n_i}{20}} * \sigma_{sens}^2 ;$$

n_t - the analyzed number of messages recorded for the stock on the day t;

σ_{sens}^2 - dispersion of the binary variable: “forecasted sign of the future return” calculated for day t.

$$\sigma_{sens}^2 = \frac{n_{t+} * n_{t-}}{n_{tsens}^2} ;$$

Where: n_{t+} - the number of messages that predict an increase of the price of the stock on day t;

n_{t-} - the number of messages that predict a decrease of the price of the stock on day t;

n_{tsens} - the number of messages that predict the sign for future returns of the specified stock during day t.

$$\frac{\sum_{i=t-19}^t n_i}{20}$$

- the average number of messages in the last twenty days prior to the analyzed day “t”.

Given the methodology of determining the indicator $I_t \in [0; 5]$, and its values are interpreted as follows:

$I_t \in [0; 0,1]$, the online activity is low, trading volume will be low, while the return sign can not be established due to high volatility.

$I_t \in (0,1;0,25]$, the online activity is moderate, trading volume will be fluctuating around the average and the price will follow the short-term trend.

$I_t \in (0,25;5]$, the online activity is high, we are in the neighborhood of an extreme point for the price trend and this indicates that the sign for the

return on short term will be opposite to the present average return. We expect a high trading volume.

Conclusions

This paper studies the impact of behavioral economics over the price and especially over the trading volume for a specific stock. The presented indicators attempt to explain this influence.

The main problem that arises in calculating these indicators is given by finding the proper vocabulary in order to determine the meaning of the messages in terms of expected sign of the future return.

The advantage of using these methods is that these alone may determine the foregoing neighborhood of a local extreme point for the share price trend.

Bibliography

- Anghelache, Gabriela; „*Piața de capital în context european*”, Editura Economică, București, 2009
- Antweiler, W.; Frank, M. „Is all that talk just noise? The information content of internet stock message boards” University of British Columbia 2001
- Ariely, Dan „*Predictably Irrational: The Hidden Forces That Shape Our Decisions*”, HarperCollins, 2008
- Das, S.; Chen, M „Yahoo! For Amazon: Opinion extraction from small talk on the web” Santa Clara University 2001
- Dellarocas, C.; Awad, N.; Zhang, X. „Using online reviews as a proxy of word-of-mouth for motion picture revenue forecasting” MIT 2004
- Felton, J.; Kim, J. „Warnings from the Enron message board” Central Michigan University, 2002
- Wysocki, Peter „Cheap talk on the web: the determinants of posting on stock message board”, University of Michigan Research 1998